



TESIS - TE142599

SKENARIO DISTRIBUSI CCTV UNTUK *SMART CITY* MENGGUNAKAN *DECISION TREE*

ARIF PRIBADI
NRP. 2215206713

DOSEN PEMBIMBING :

1. Mochamad Hariadi, S.T., M.Sc., Ph.D.
2. Dr. Supeno Mardi Susiki Nugroho, S.T., M.T.

**PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA /
CHIEF INFORMATION OFFICER (CIO)
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**



TESIS - TE142599

SKENARIO DISTRIBUSI CCTV UNTUK *SMART CITY* MENGGUNAKAN *DECISION TREE*

ARIF PRIBADI
NRP. 2215206713

DOSEN PEMBIMBING :

1. Mochamad Hariadi, S.T., M.Sc., Ph.D.
2. Dr. Supeno Mardi Susiki Nugroho, S.T., M.T

**PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA /
CHIEF INFORMATION OFFICER (CIO)
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**

LEMBAR PENGESAHAN

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Teknik (M.T)
di
Institut Teknologi Sepuluh Nopember
oleh:

Arif Pribadi
NRP. 2215206713

Tanggal Ujian : 6 Januari 2017
Periode Wisuda : Maret 2017

Disetujui oleh:

1. **Mochamad Hariadi, S.T., M.Sc., Ph.D.** (Pembimbing I)
NIP. 19691209 199703 1 002
2. **Dr. Supeno Mardj Susiki Nugroho, S.T., M.T.** (Pembimbing II)
NIP. 19700313 199512 1 001
3. **Dr. Eko Mulyanto Yuniarno, S.T., M.T.** (Penguji)
NIP. 19680601 199512 1 009
4. **Dr. Istas Pratomo, S.T., M.T.** (Penguji)
NIP. 19790325 200312 1 001
5. **Dr. Diah Puspito Wulandari, S.T., M.Sc.** (Penguji)
NIP. 19801219 200501 2 001

an. Direktur Program Pascasarjana

Asisten Direktur

Prof. Dr. Ir. Iri Widjaja, M.Eng.
NIP. 19611021 198603 1 001



Direktur Program Pascasarjana

Prof. Ir. Djauhar Manfaat, M.Sc, Ph.D
NIP. 19601202 198701 1 001

Halaman ini sengaja dikosongkan

PERNYATAAN KEASLIAN TESIS

Dengan ini saya menyatakan bahwa isi keseluruhan Tesis saya dengan judul “**Skenario Distribusi CCTV untuk *Smart City* menggunakan *Decision Tree***” adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diijinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri.

Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, Januari 2017

Arif Pribadi
NRP. 2215206713

Halaman ini sengaja dikosongkan

SKENARIO DISTRIBUSI CCTV UNTUK *SMART CITY* MENGUNAKAN *DECISION TREE*

Nama Mahasiswa : Arif Pribadi
NRP : 2215206713
Dosen Pembimbing I : Mochamad Hariadi, S.T., M.Sc., Ph.D.
Dosen Pembimbing II : Dr. Supeno Mardi Susiki Nugroho, S.T., M.T.

ABSTRAK

Smart City telah berkembang ke beberapa negara, termasuk Indonesia. Dengan *Smart City* kebutuhan penduduk atas kesehatan, pendidikan, transportasi, komunikasi, keamanan akan mudah terwujud. CCTV sebagai salah satu pendukung *Smart City* banyak dipasang di berbagai sudut kota. Tujuannya agar pemerintah dapat memantau kondisi lingkungan. Perempatan, Pertigaan, bundaran merupakan daerah rawan kemacetan yang biasa dipantau oleh CCTV. Begitu juga dengan jalan sepi dan gelap berpotensi rawan tindak kejahatan. Pemilihan jenis CCTV akan berbeda untuk masing-masing daerah. misalnya daerah gelap akan diberikan dengan CCTV Infrared.

Penelitian ini akan menerapkan teknologi klasifikasi untuk membangun model prediksi. Model Prediksi akan digunakan untuk pemilihan CCTV yang tepat. Penelitian ini menggunakan 108 sampel yang telah dipasang di sebuah kota. Dengan algoritma *decision tree* (C4.5) diperoleh tingkat akurasi 89,81% model prediksi.

Kata Kunci : *Decision Tree* , *CCTV*, *Smart City*.

Halaman ini sengaja dikosongkan

SCENARIO FOR DISTRIBUTION CCTV SMART CITY USING DECISION *TREE*

By : Arif Pribadi
Student Identity Number : 2215206713
Supervisor (s) : 1. Mochamad Hariadi, S.T., M.Sc., Ph.D.
2. Dr. Supeno Mardi Susiki Nugroho, S.T., M.T.

ABSTRACT

Smart City has grown to several countries, including Indonesia. Smart City brings hope to ensure the welfare of the population. Smart City with the needs of the population's health, education, transport, communications, security will be easily realized. CCTV as one of the supporting Smart City widely installed in various corners of the city. The goal for the government is to monitor environmental conditions. Intersection, T-junction, the roundabout is an area prone to congestion commonly monitored by CCTV. as well as residents and dark lonely road it will be prone to crime. Selection of the type of CCTV is different for each region. eg dark areas will be given with Infrared CCTV.

This study will apply the classification technology to build predictive models. This prediction of the model will be used for the selection of the right CCTV. This study using 108 samples that have been installed in a city. With Decision *Tree* algorithm obtained accuracy rate of 89,81% prediction models.

Keyword : *Decision Tree , CCTV, Smart City.*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Alhamdulillahirobbil'alamin, puji syukur atas segala limpahan nikmat dan karunia Allah SWT, Tuhan yang Maha Kuasa. Hanya dengan petunjuk, rahmat dan ridho-Nya, sehingga penulis dapat menyelesaikan tesis dengan judul **“Klasifikasi Distribusi CCTV untuk *Smart City* menggunakan *Decision Tree*”**.

Ucapan terima kasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya saya sampaikan kepada yang terhormat Mochamad Hariadi, S.T., M.Sc., Ph.D. selaku pembimbing pertama dan Dr. Supeno Mardi Susiki Nugroho, S.T., M.T. selaku pembimbing kedua, yang dengan penuh perhatian, dan kesabaran selalu meluangkan waktu, memberikan pengarahan dan motivasi serta semangat dalam penulisan tesis ini.

Penulis dapat menyelesaikan tesis ini, juga tidak terlepas dari bantuan dan kerjasama dari berbagai pihak, maka perkenankan saya dengan sepenuh hati menyampaikan terima kasih yang tak terhingga kepada:

1. Prof. Ir. Joni Hermana, M.Sc.Es, Ph.D., selaku Rektor Institut Teknologi Sepuluh Nopember Surabaya, yang telah memberikan kesempatan dan fasilitas kepada saya untuk mengikuti dan menyelesaikan pendidikan pada Program Studi Magister Keperawatan.
2. Kementerian Komunikasi dan Informasi yang telah memberikan kesempatan mendapatkan beasiswa Program Magister (S2) Telematika/Chief Information Officer pada Institut Teknologi Sepuluh Nopember Surabaya.
3. Dr. Bambang Lelono Widjiantoro, S.T., M.T., selaku Dekan Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember Surabaya yang telah memberikan arahan, kesempatan dan fasilitas kepada kami untuk mengikuti dan menyelesaikan pendidikan di Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember Surabaya.
4. Dr. Eng. Ardyono Priyadi, ST, M.Eng., selaku Ketua Jurusan Teknik Elektro, Fakultas Teknologi Industri Institut Teknologi Sepuluh Nopember Surabaya yang telah memberikan kesempatan dan fasilitas kepada kami untuk

mengikuti pendidikan pada Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember Surabaya.

5. Dr. Ir. Djoko Purwanto, M.Eng, selaku Kepala Program Studi Magister (S2) dan Doktor (S3), Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember Surabaya yang telah memberikan kesempatan dan fasilitas kepada kami untuk mengikuti pendidikan pada Program Magister (S2) Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember Surabaya.
6. Dr. Adhi Dharma Wibawa, ST, MT, selaku Koordinator Bidang Keahlian Telematika / Chief Information Officer (CIO) sekaligus Dosen Pembimbing Akademik CIO 2015 Jurusan Teknik Elektro, atas arahan, bimbingan dan motivasinya dalam menyelesaikan perkuliahan maupun penulisan tesis ini.
7. Seluruh Pengajar dan staf Program Studi Magister (S2) Jurusan Teknik Elektro, Bidang Keahlian Telematika/Chief Information Officer (CIO), yang telah mentransfer ilmu pengetahuannya melalui kegiatan perkuliahan maupun praktikum serta membantu kelancaran pengurusan administrasi perkuliahan dan penyelesaian tesis ini.
8. Muhammad Afib, S. Sos selaku Kepala Dinas Sosial Tenaga Kerja dan Transmigrasi Kabupaten Pekalongan, atas segala arahan dan bimbingan dalam menyelesaikan pendidikan ini.
9. Edi Apriyanto, S.H selaku Kepala Balai Latihan Kerja, Dinas Sosial Tenaga Kerja dan Transmigrasi Kabupaten Pekalongan yang telah memberikan ijin dan kesempatan kepada penulis untuk melanjutkan pendidikan ke jenjang yang lebih tinggi.
10. Rekan-rekan pengajar di Balai Latihan Kerja, Dinas Sosial Tenaga Kerja dan Transmigrasi Kabupaten Pekalongan yang selalu mendoakan dan memberi dorongan semangat sehingga penulis bisa menyelesaikan pendidikan dan penulisan tesis ini tepat waktu.
11. Orangtuaku Amat Alfian (Alm) dan Hj. Sri Iriani serta H. Aslimin dan Hj. Siti Cholasoh terimakasih atas segala do'a dan dukungannya sehingga penulis dapat menyelesaikan tesis ini tepat waktu.

12. Istriku tercinta drg. Endra Aryani dan anak-anakku tersayang Aqila Pratama Gumilang dan Zafar Gigih Prawira, yang dengan penuh kesabaran, cinta, kasih-sayang dan do'a yang tiada henti dipanjatkan demi selesainya studi ini. Semoga keluarga kita selalu mendapat ridlo-Nya dan menjadi keluarga yang sakinah mawaddah warohmah. Amin.
13. Saudara-saudaraku tersayang Mila, Ryan, Opi, Vera, Ato, Endang, Ita, Uut, Habib, Umi, Ida dan seluruh keponakanku, yang banyak membantu do'a.
14. Mahasiswa Program Studi Magister Telematika / Chief Information Officer (CIO) Angkatan 2015 yang selalu kompak dan saling mendukung, saling mendoakan baik dalam perkuliahan maupun dalam penyelesaian penulisan tesis ini.
15. Mahasiswa Program Studi Magister Telematika / Chief Information Officer (CIO) Angkatan 2014 khususnya tim 7 (fadli, ulfa, pram, narti, palti, ricki, dan delphi) yang saling mendukung.

Semoga Allah SWT membalas kebaikan semua pihak yang telah memberi kesempatan, dukungan dan bantuan dalam menyelesaikan tesis ini. Penulis menyadari bahwa tesis ini masih jauh dari sempurna, oleh karena itu saran dan kritik yang membangun sangat diharapkan demi kesempurnaan tulisan ini, sehingga tesis ini memberikan manfaat yang baik bagi agama, bangsa dan negara.

Surabaya, Januari 2017

Penulis

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	iii
ABSTRAK	vii
KATA PENGANTAR	xi
DAFTAR ISI	xv
DAFTAR GAMBAR	xvii
DAFTAR TABEL	xix
BAB I PENDAHULUAN	1
1.1. Latar belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	3
1.4. Manfaat	3
1.5. Batasan Masalah	3
BAB II KAJIAN PUSTAKA DAN DASAR TEORI	5
2.1. Kajian Pustaka	5
2.2. Landasan Teori	7
2.2.1. <i>Smart City</i>	7
2.2.2. Kamera CCTV	8
2.2.3 Data Mining	14
2.2.3.1 Cross-validation	16
2.2.3.2 Evaluasi	17
2.2.3.3 Algoritma <i>Decision Tree</i>	18
2.2.3.4 Pruning	22
BAB III METODELOGI PENELITIAN	25
3.1 Tahap Pre-processing Data	26
3.2 Validasi	26
3.3 <i>Decision tree</i>	27
3.4 Evaluasi	38

BAB IV ANALISA DAN PEMBAHASAN	41
4.1 Analisa Data Awal	41
4.2 Analisa Performa Tipe CCTV	42
4.3 <i>Klasifikasi Decision Tree</i>	44
BAB V KESIMPULAN	59
DAFTAR PUSTAKA.....	xxi

DAFTAR GAMBAR

Gambar 2.1 Kamera Bullet/Fixed	10
Gambar 2.2 Kamera PTZ	10
Gambar 2.3 Perbandingan 4CIF IP Camera 4 CIF Analog Camera	11
Gambar 2.4 Perbandingan CIF dengan 1,3MP	11
Gambar 2.5 Perbandingan foto dengan lensa 3.6 – 12mm	13
Gambar 2.6 Perbandingan lensa mm dan sudut (angle field of view)	13
Gambar 2.7 Definisi klasifikasi.....	15
Gambar 2.8 Pendekatan umum proses klasifikasi.....	15
Gambar 2.10 <i>Two-fold-cross-validation</i>	16
Gambar 2.12 <i>k-cross validation</i>	17
Gambar 2.13 <i>Binary Attribute</i>	19
Gambar 2.14 <i>Nominal Attribute</i>	20
Gambar 2.15 <i>Ordinal Attribute</i>	20
Gambar 2.6 <i>Contionuos Attribute</i>	21
Gambar 3.1 Metodologi Penelitian	25
Gambar 3.2 Ilustrasi <i>10-fold validation</i>	26
Gambar 3.3 <i>Induction Tree</i>	27
Gambar 3.4 <i>Tree</i> Hasil Tahap 1	31
Gambar 3.5 <i>Tree</i> Hasil Tahap 2	34
Gambar 3.6 <i>Tree</i> Hasil Tahap 3	37
Gambar 3.7 Flow Chart Data Keputusan Bermain Tennis	38
Gambar 4.1 Rule ke-1 Dataset Kamera.....	45
Gambar 4.2 Rule Ke-2 Dataset Kamera.....	46
Gambar 4.3 Rule Ke-3 Dataset Kamera.....	47
Gambar 4.4 Rule Ke-4 Dataset Kamera.....	48
Gambar 4.5 Rule Ke-5 Dataset Kamera.....	49
Gambar 4.6 Rule Ke-6 Dataset Kamera.....	50
Gambar 4.7 Rule Ke-7 Dataset Kamera.....	51
Gambar 4.8 Rule Ke-8 Dataset Kamera.....	52

Gambar 4.9 Rule Ke-9 Dataset Kamera	53
Gambar 4.10 Rule Ke-10 Dataset Kamera	54
Gambar 4.11 Klasifikasi benar Kelompok Rule ke-1.....	56
Gambar 4.12 Klasifikasi benar Kelompok Rule ke-2.....	56
Gambar 4.13 Klasifikasi benar Kelompok Rule ke-3.....	57

DAFTAR TABEL

Tabel 2.1	Standard Resolusi CCTV	12
Tabel 2.2	Confusion Matrix	17
Tabel 3.1	Data Keputusan Permainan Tennis	28
Tabel 3.2	Hasil Perhitungan Entropi Cuaca Tahap 1	30
Tabel 3.3	Hasil Perhitungan Gain Ratio Tahap 1	31
Tabel 3.4	Data Keputusan Permainan Tennis Tahap 2	32
Tabel 3.5	Hasil Perhitungan Entropi Kelembaban Tahap 2	33
Tabel 3.6	Hasil Perhitungan Gain Ratio Tahap 2	34
Tabel 3.7	Data Keputusan Permainan Tennis Tahap 3	35
Tabel 3.8	Hasil Perhitungan Entropi Kelembaban Tahap 3	36
Tabel 3.9	Hasil Perhitungan Gain Ratio Tahap 3	37
Tabel 3.10	Confusion Matrix Data Keputusan Permainan Tennis	39
Tabel 4.1	Dataset Kamera CCTV	41
Tabel 4.2	Label Dataset Kamera	42
Tabel 4.3	Spesifikasi CCTV Tipe A, Tipe B, dan Tipe C	43
Tabel 4.4	Klasifikasi Benar 10 Rule Decision <i>Tree</i>	55

Halaman ini sengaja dikosongkan

BAB I

PENDAHULUAN

1.1. Latar belakang

Tindak kejahatan semakin tahun semakin tinggi. Hal ini berbanding lurus dengan angka pertumbuhan penduduk dan berbanding terbalik dengan tersedianya lapangan pekerjaan. Dilain pihak teknologi semakin berkembang, bagi yang tidak mengikuti perkembangan teknologi akan tertinggal dari yang lain. Karena dengan teknologi, manusia semakin mudah dalam menjalani kehidupan. Kemudahan-kemudahan yang tersedia ini merupakan salah satu dari perwujudan *Smart City*. Dimana *Smart City* bertujuan untuk menciptakan suatu entitas yang terintegrasi secara rapi yang mempermudah manusia dalam kehidupannya.

Kamera CCTV merupakan salah satu contoh dari aplikasi *Smart City*, dengan kamera ini maka pihak pemerintah dapat mengawasi lingkungan secara real time selama 24 jam. Adanya pengawasan ini membuat rasa aman pada masyarakat.

Saat ini beberapa kota mulai berlomba dalam menerapkan *Smart City*. Atau biasa disebut Kota Cerdas. Pemimpin daerah seakan mengambil kesimpulan kesuksesan daerah yang dipimpinnya dinilai dari peringkat penerapan *Smart City*. Berbekal anggaran dana yang besar, daerah mulai membangun *Smart City* secara bertahap. Di Indonesia ada beberapa kota yang telah menerapkan *Smart City*, antara lain Jakarta, Surabaya, Bandung dan Makasar. Di kota Bandung, pemerintah kota Bandung memiliki Room Monitoring yang dilengkapi dengan TV layar datar berukuran besar. Dari gambar yang terlihat di Room Monitoring, pemerintah kota Bandung nantinya akan mengeluarkan kebijakan untuk mengatur Kota. Agar kebijakan yang diambil nanti sesuai dengan kondisi riil yang terjadi, maka pemilihan kamera menjadi penting.

Saat ini kota mengalami krisis, karena permintaan yang melebihi pasokan. Misalnya air, energi (listrik), transportasi, kesehatan, pendidikan dan keamanan. Untuk meningkatkan efisiensi penggunaan anggaran dan kualitas hidup penduduk,

maka kota mengembangkan Teknologi Informasi dan Komunikasi. Penerapan TIK inilah yang mengawali berkembangnya *Smart City* (Naphade dkk, 2011).

Inti dari perkembangan kota menjadi *Smart City* adalah untuk pengembangan ekonomi. Pada pertengahan tahun 2008-2009 saat dunia menghadapi krisis ekonomi, kota menyadari bahwa persaingan antar kota tidak hanya dengan kota tetangga. Namun dengan perkembangan jaringan internet, pesaing datang dari kota-kota lain di penjuru dunia. Persaingan tidak hanya pada investasi dan pekerjaan, namun juga bersaing untuk Generasi Y dan Z yang diharapkan akan membawa kekuatan ekonomi. Untuk itu maka kota menyiapkan lingkungan yang menarik kedatangan Generasi Y dan Z. (Harrison dkk, 2011). Terbukti bahwa *Smart City* tumbuh didukung banyaknya penduduk yang berpendidikan tinggi di kota tersebut. (Winters, 2011)

Salah satu penelitian yang dilakukan di kota Gwang Myeong Provinsi Gyeonggi Korea Selatan menyimpulkan bahwa setelah terpasang kamera CCTV, angka kriminalitas (perampokan dan pencurian) pada daerah tersebut mengalami penurunan sebesar 47,4%. Dan daerah yang tidak terpasang kamera CCTV tidak mengalami perubahan angka. Daerah yang bertetangga dengan kota tersebut juga mengalami penurunan angka kriminalitas. (Park dkk, 2012)

Penerapan CCTV yang tepat terbukti dapat mengurangi angka kriminalitas pada kota Gwang Myeong di Korea Selatan. Dalam penelitian kali ini akan diteliti skenario pemetaan kamera CCTV yang tepat agar kebijakan pemerintah yang akan dikeluarkan mewakili kondisi masyarakat sesungguhnya.

1.2. Rumusan Masalah

1. Pemasangan kamera CCTV tidak tepat sesuai kondisi permasalahan di lokasi.
2. Pemilihan tipe kamera CCTV yang tidak sesuai dengan peruntukannya.

1.3. Tujuan

1. Memonitor kondisi jalan sehingga dihasilkan rekayasa lalu lintas yang tepat untuk mengurangi kecelakaan, kriminalitas dan kemacetan di jalan raya.
2. Mendapatkan model prediksi distribusi CCTV.

1.4. Manfaat

Memberikan solusi bagi pemerintah agar dapat menempatkan CCTV sesuai kondisi lokasi sesuai rule yang dihasilkan penelitian ini.

1.5. Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Lokasi pengambilan sampel adalah dikota Makasar.
2. CCTV yang diamati hanya yang mengawasi jalan raya.
3. Pengambilan sampel dilakukan sekali pada tahun 2015.
4. Parameter yang digunakan sebagai pertimbangan pemetaan yaitu; Rawan Kriminal, Rawan Kecelakaan, Rawan Kemacetan, Temu Lalu Lintas, dan Jenis jalan (Lurus, Pertigaan, Perempatan, Bundaran dan Lorong).
5. Variabel harga tidak diikutsertakan dalam penelitian ini.

Halaman ini sengaja dikosongkan

BAB II

KAJIAN PUSTAKA DAN DASAR TEORI

2.1. Kajian Pustaka

Penelitian tentang *Smart City* terus berkembang. Kebutuhan manusia akan kenyamanan, untuk membuat segala yang memungkinkan dirubah dari konvensional menjadi berbasis *Information Communication and Technology* (ICT). Misalnya, dahulu untuk menaiki kendaraan umum harus langsung ke tepi jalan untuk menghentikan kendaraan, namun saat ini cukup dengan aplikasi transportasi maka dapat dengan mudah memanggil penyedia jasa transportasi. Banyaknya penyedia jasa layanan transportasi memunculkan penelitian tentang perilaku dan permintaan traveler. Dengan bantuan GPS dapat diketahui moda transportasi yang tepat dari posisi traveler sehingga waktu dan biaya yang diperlukan berkurang secara signifikan. Penelitian tersebut mengklasifikasi data GPS dari beberapa mode transportasi (mobil, jalan kaki, sepeda, kereta api dan bus) dengan algoritma SVM. (Bolbol dkk, 2012)

Penelitian lain mengenai *Smart City* yang digunakan untuk mengurangi kemacetan selama perjalanan karena meningkatnya jumlah kendaraan pribadi. Dengan mengetahui pola perjalanan, penyedia layanan transportasi diharapkan untuk menyesuaikan strategi pemasaran dan memberikan kepuasan pelanggan. Penelitian dilakukan di Beijing China untuk mengevaluasi rute perjalanan tourist menggunakan Smart Card. Data tersebut kemudian diklasifikasikan oleh *K-Means* dan teori *Rough-Set*. (Maa dkk, 2013)

Bila pada penelitian tersebut diatas digunakan data GPS, maka dalam penelitian yang lain digunakan data CCTV untuk mengevaluasi kinerja *Smart City*. CCTV sering digunakan untuk mengidentifikasi pelaku tindak kejahatan. Salah satu analisa yang digunakan adalah *forensic gait* atau analisa berdasarkan cara berjalan seseorang. Analisa jenis ini sering terkendala dengan *frame rate* rekaman dari CCTV. Frame rate biasa dinyatakan dalam fps (frame per secon), atau jumlah potongan gambar dalam satu detik. Frame rate berbanding lurus dengan nilai

identifikasi *forensic gait*. Semakin besar frame rate rekaman CCTV, semakin memudahkan proses identifikasi *forensic gait*. (Birch dkk, 2014). Penelitian serupa juga digunakan untuk melacak orang dan kendaraan dengan proyeksi 3D CCTV (Edelman dan Bijhold 2010). Penelitian tentang faktor-faktor penyebab kecelakaan lalu lintas pernah dilaksanakan di kota Leeds Inggris. Penelitian ini menitik beratkan pada alasan dan frekuensi terjadinya kecelakaan yang kemudian diklasifikasikan menjadi faktor penyebab kecelakaan. Alasan penggunaan CCTV karena CCTV memberikan suatu pandangan yang independent, sehingga keputusan yang diambil pemerintah untuk mencegah terjadinya kecelakaan dapat diputuskan dengan tepat. (Conche dkk, 2006). Penelitian lebih lanjut tentang CCTV terkait dengan kecepatan kendaraan. Kecepatan dihitung berdasarkan panjang kendaraan dikaitkan dengan 2 frame rekaman CCTV dengan metode *invariant geometri proyektif cross-ratio*. Kalibrasi Doppler Radar digunakan untuk memvalidasi kecepatan tersebut. (Wong dkk, 2014). Selain itu, CCTV juga telah dikembangkan di Thailand untuk mendukung World Heritage City (WHC). CCTV digunakan untuk mengawasi area dan menghitung jumlah pengunjung masuk dan keluar objek wisata di Thailand. Sistem ini memiliki akurasi 80-90% untuk menghitung jumlah pengunjung. Sebagian besar kesalahan terjadi ketika menghitung daerah dengan latar belakang gelap atau ketika menghitung orang yang berjalan dalam kelompok. (Nerngchamnong dkk, 2011)

Secara umum rekaman CCTV disimpan dalam bentuk enkripsi di kartu memori. Ketika terjadi kejahatan, rekaman tersebut diambil kemudian dienkripsi dan dilihat oleh pihak yang berwenang. Bila tidak terjadi kejahatan, maka rekaman video ditimpa oleh rekaman yang baru setelah beberapa minggu secara otomatis. Hal ini melindungi privasi masyarakat. (Prashyanusorn dkk, 2011)

Pada tahun 2015 telah dilakukan penelitian tentang pemodelan distribusi hotspot pada kota-kota yang dianggap maju. Penelitian tersebut melibatkan 10 kota besar di Eropa dengan ribuan hotspot (Michael dkk, 2015). Pemodelan ini dilakukan untuk mengurangi beban *offloading* dari jaringan GSM. Penelitian ini mengklasifikasikan hubungan antara jumlah penduduk dengan jumlah hotspot Wi-Fi dengan algoritma *Fuzzy C-Means*.

Penelitian ini akan mencari model prediksi dari kamera CCTV yang telah terpasang di kota makasar, kemudian dari model prediksi tersebut dihitung nilai akurasi insatalasi CCTV pada kota Makasar.

2.2. Landasan Teori

2.2.1. *Smart City*

Semakin berkembangnya teknologi informasi, banyak kota-kota didunia yang mulai mencanangkan konsep tata kota yang serba modern. *Smart City* menjadi indikator baru untuk menentukan sebuah kota layak atau tidak untuk dihuni. Saat ini, kebutuhan dasar manusia (kebutuhan fisiologis, keamanan, cinta, harga diri, dan aktualisasi diri telah dipengaruhi oleh faktor-faktor seperti globalisasi, teknologi, dan pertumbuhan penduduk perkotaan. Hak untuk hidup di lingkungan yang bersih dan hijau, untuk menerima perawatan kesehatan dan sosial secara efisien dan tepat waktu, dan dapat dengan mudah dihubungi, terhubung, dan berkomunikasi adalah beberapa contoh dari kehidupan modern, Hal ini menjadi sebuah tantangan yang harus dipecahkan oleh pemerintah, terutama di lingkungan perkotaan. Untuk itulah maka diterapkanlah konsep *Smart City*.

Dalam mengembangkan *Smart City*, segala sesuatu dalam kota termasuk orang, bisnis, teknologi, proses, data, infrastruktur, konsumsi, ruang, energi, strategi, manajemen, dan sebagainya harus menjadi pintar. Mengingat semuanya pintar berarti komponen ini harus terhubung dan saling mendukung, Oleh karena itu, tidak hanya teknologi cerdas dan inovatif yang dibutuhkan, namun juga teknologi yang terintegrasi dengan baik dalam sebuah sistem

Sebuah lembaga konsultan global, *Price Waterhouse Cooper (PWC)* memaparkan ada lima kriteria keterwujudan predikat *Smart City* pada sebuah kota, antara lain: memiliki sistem transportasi terintegrasi, memanfaatkan teknologi untuk meningkatkan rasa keamanan warganya, melakukan manajemen lingkungan yang efektif, memiliki sentra inovasi, serta mampu memberikan layanan *e-government*

Di Indonesia sendiri, kriteria *Smart City* juga pernah diungkapkan Menteri Pekerjaan Umum dan Perumahan Rakyat (PURR) Basuki Hadimoeljono. delapan acuan utama yang bisa memperkuat gagasan *Smart City*, yaitu *smart development planning*, *smart green open space*, *smart transportation*, *smart waste management*, *smart water management*, *smart building*, dan *smart energy*.

2.2.2. Kamera CCTV

Istilah kamera berasal dari bahasa latin “camera obscura” yang berarti ruang gelap. Ruang ini merupakan alat bantu seniman selama Abad Pertengahan. Seniman menggunakan ruang *lightproof* (blok cahaya), dalam bentuk kotak, dengan lensa cembung di satu ujung dan layar yang bereaksi gambar di lain, untuk melacak gambar dan kemudian menghasilkan lukisan.

Penelitian pertama tentang kamera televisi dilakukan pada tahun 1930 oleh insinyur Rusia Vladimir Zworykin (1889-1982). Kamera pertama dibuat pada tahun 1931, fokus gambar diarahkan ke sebuah mosaik sel fotolistrik. Tegangan induksi di setiap sel adalah ukuran dari intensitas cahaya dan dapat ditransmisikan sebagai sinyal listrik. Konsep ini dan beberapa modifikasi kecil tetap bertahan selama beberapa dekade. Kamera pertama tersebut dibuat dengan tabung kaca dan lapisan fosfor yang peka terhadap cahaya dalam kaca. Kamera jenis ini kemudian disebut kamera tabung. (Damjanovski, 2014)

Untuk memahami CCTV tidaklah sulit. Pada fotografi yang terdahulu, informasi optik (gambar) dikonversi menjadi jejak emulsi kimia. Di kamera televisi, kita mengubah informasi optik menjadi sinyal listrik. Semua menggunakan lensa dengan panjang fokus tertentu dan sudut pandang tertentu, namun berbeda untuk format yang berbeda. Saat ini fotografi telah benar-benar berubah. Semuanya telah berubah menjadi digital. Kelebihan digital ada pada detail dan kualitas.

CCTV pertama kali dipasang oleh Siemens AG pada saat Test Stand VII di Peenemünde Jerman pada tahun 1942, untuk mengamati peluncuran roket V-2. Walter Bruch, Wayne Cox, dan Tashara Arnold bertanggung jawab untuk desain teknologi dan instalasi sistem CCTV tersebut.

CCTV (Closed Circuit Television) merupakan alat yang digunakan untuk merekam kejadian di area tertentu untuk ditampilkan di monitor dengan cakupan publikasi yang terbatas. Di kota yang menganut konsep *Smart City*, nantinya kamera CCTV bisa berperan lebih aktif dan mampu melakukan pengawasan secara mandiri. Data dari CCTV dan perangkat-perangkat lain kemudian dihubungkan ke pusat *server*.

Namun CCTV tidak bisa terpasang di semua tempat, seperti yang tercantum dalam daftar persyaratan hukum yang harus dipenuhi oleh sistem CCTV berikut ini (Kurdi, 2013)

- CCTV hanya boleh digunakan di tempat umum, yaitu di daerah di mana orang tidak memiliki ekspektasi privasi yang tinggi.
- Persetujuan harus diperoleh dari pemilik bangunan dan struktur kamera harus dilampirkan.
- Teknisi dipilih secara hati-hati untuk memastikan profesionalisme dan kepercayaan karena merekam materi sensitif dan rahasia.
- Materi yang direkam harus diperlakukan sangat rahasia dan dihapus setelah jangka waktu tertentu atau tidak lebih dari 30 hari, kecuali diperlukan untuk bukti. Hasil rekaman harus tetap terkunci dan teregister untuk mengontrol penyimpanan rekaman,
- Harus ada sanksi yang melekat bagi pelanggaran pedoman ini.

Teknologi kamera CCTV dapat di labelkan sebagai berikut :

1. Kamera Biasa, hanya menangkap gambar sesuai dengan yang di terima oleh CMOS (sensor kamera yang berfungsi menangkap gambar)
2. Thermal kamera, berfungsi untuk mendapatkan gambar dari suhu object.
3. Infra Red Kamera, berfungsi untuk mendapatkan object dari ruangan yang sangat gelap.

Dari sisi label bentuk CCTV dapat dibagi menjadi 2 macam,

1. Bullet / fixed (posisi kamera Θ tidak berubah-ubah)

Yaitu jenis kamera yang objek sasarannya tidak dapat diubah. Contoh CCTV tipe bullet dapat dilihat pada Gambar 2.1.



Gambar 2.1 Kamera Bullet/Fixed

2. PTZ (Pain, tilt, Zoom)

yaitu kamera yang dapat digerakkan ke kiri dan ke kanan juga kebawah dan keatas serta memiliki kemampuan untuk *zoom* sasaran object dengan kelipatan berkali-kali. Contoh CCTV tipe PTZ dapat dilihat pada Gambar 2.2.



Gambar 2.2 Kamera PTZ

Perbedaan CCTV Analog dengan IP Camera

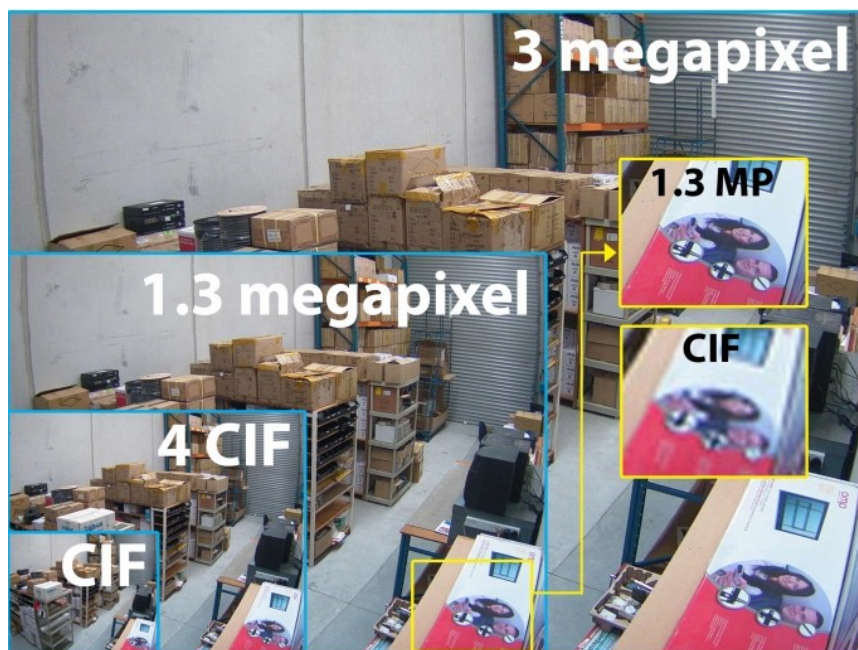
Kamera Standard Definition (SD) , High Definition (HD) dan IP Camera dapat dibedakan dari resolusi (kualitas dan ukuran) gambar yang direkam. Resolusi maksimal Kamera CCTV analog terbatas sampai D1 720x576 piksel (PAL) atau 0,4 megapixel. Bahkan jika kamera dengan resolusi lebih tinggi (700 TVL = 1020X596 piksel) digunakan maka hanya akan merekam pada resolusi D1. Kualitas video yang direkam dengan sistem analog berkurang karena proses konversi gambar digital ke analog dan kemudian dikonversi kembali ke digital untuk ditampilkan. Sementara pada IP Camera gambar langsung direkam ke dalam Network Video Recorder dalam bentuk digital. Hal ini yang membedakan

CCTV berbasis IP Camera dengan sistem analog. Perbandingan IP Camera dengan Analog dengan kualitas sama 4CIF dapat dilihat pada Gambar 2.3.



Gambar 2.3 Perbandingan 4CIF IP Camera 4 CIF Analog Camera

Selanjutnya untuk mengetahui perbedaan dari kamera kualitas CIF dengan kamera kualitas 1,3MP dapat dilihat pada Gambar 2.4. (Wireless CCTV 2016). Hasil rekaman menggunakan kamera CIF dan 1,3 MP memiliki perbedaan yang signifikan ketika gambar diperbesar.



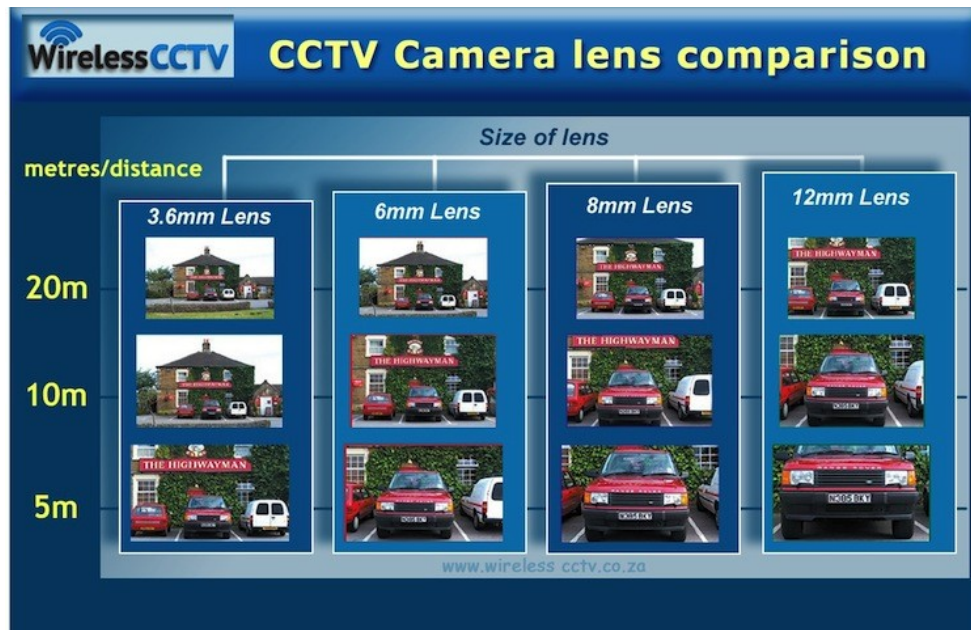
Gambar 2.4 Perbandingan CIF dengan 1,3MP

Kualitas kamera CCTV menentukan hasil pengolahan data, misalnya pada saat CCTV digunakan untuk pencatatan nomer kendaraan bermotor. Jika kualitas rendah maka nomor kendaraan tidak dapat terbaca. Kualitas ketajaman kamera CCTV sering dinyatakan dalam piksel. Kamera kualitas CIF mempunyai ketajaman gambar 704 x 240 piksel. Sedangkan kamera kualitas 1,3 MP mempunyai ketajaman 1280 x 1024 piksel. Standard resolusi kamera CCTV untuk tipe yang lain dapat dilihat pada Tabel 2.1.

Tabel 2.1 Standard Resolusi CCTV

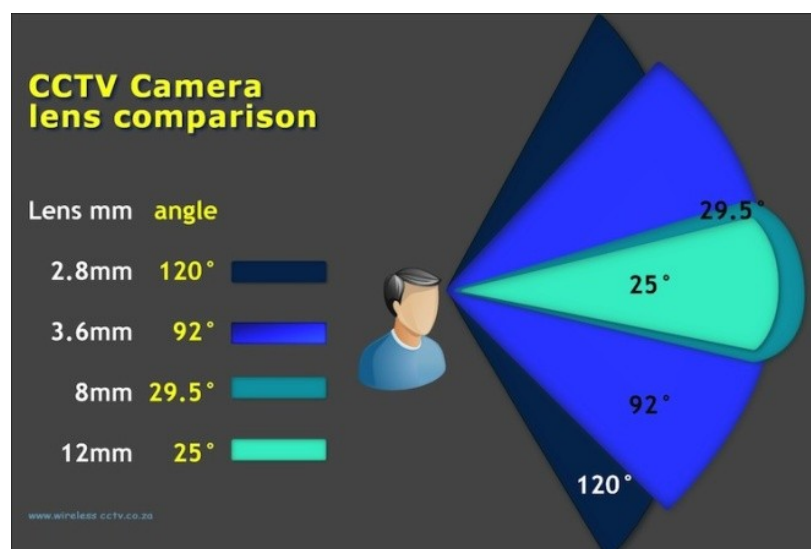
Term	Pixels (W x H)	Notes
QCIF	176 x 120	Quarter CIF (half the height and width as CIF)
CIF	352 x 240	
2CIF	704 x 240	2 times CIF width
4CIF	704 x 480	2 times CIF width and 2 times CIF height
D1	720 x 480	aka "Full D1"
720p HD	1280 x 720	720p High Definition
960p HD	1280 x 960	960p High Definition (Sony HD standard)
1.3 MP	1280 x 1024	aka "1 Megapixel" or "1MP"
2 MP	1600 x 1200	2 Megapixel
1080p HD	1920 x 1080	1080p High Definition
3 MP	2048 x 1536	3 Megapixel
5 MP	2592 x 1944	5 Megapixel

Selain dibedakan dari resolusi kamera, kualitas CCTV juga sering dibedakan dari ukuran lensa. Lensa mempengaruhi kualitas objek ketika diperbesar atau diperkecil (zoom out/zoom in). Dengan jarak yang sama saat perekaman obyek, lensa dengan ukuran lebih besar dapat menangkap obyek dengan lebih jelas (detil). Apabila obyek direkam dalam jarak yang berbeda, maka obyek yang direkam dalam jarak yang lebih dekat akan menghasikan kualitas gambar yang lebih jelas. Ilustrasi tentang fungsi lensa pada kamera dapat dilihat pada Gambar 2.5 (Wireless CCTV 2016).



Gambar 2.5 Perbandingan foto dengan lensa 3.6 – 12mm

Sedangkan angel field of view menunjukkan besarnya jangkauan kamera. Semakin besar sudut, maka makin lebar jangkauan kamera, namun jarak obyek yang dapat diamati semakin pendek. Sebaliknya, semakin kecil sudut maka semakin sempit jangkauan kamera, namun jarak obyek yang dapat diamati semakin jauh. Ilustrasi fungsi angel field of view dapat dilihat pada Gambar 2.6



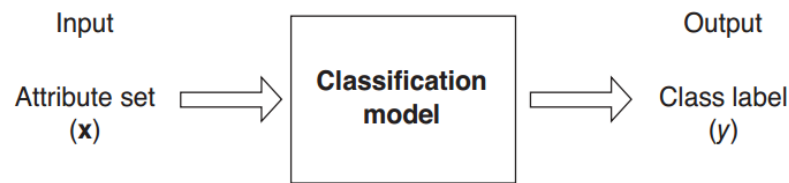
Gambar 2. 6 Perbandingan lensa mm dan sudut (angle field of view)

2.2.3 Data Mining

Kehidupan di dunia saat ini penuh dengan data. Setiap hari, orang menjumpai informasi dalam jumlah besar dan menyimpannya sebagai data, dan biasa digunakan untuk analisis lebih lanjut dan manajerial. Salah satu cara vital dalam menangani data ini adalah dengan mengklasifikasikan atau mengelompokkan data ke dalam satu set label atau kluster. Klasifikasi, memainkan peran penting dan sangat diperlukan dalam sejarah panjang perkembangan manusia. Dalam rangka untuk mempelajari sesuatu atau memahami fenomena baru, orang selalu mencoba untuk mencari ciri yang dapat menggambarkannya, dan selanjutnya membandingkannya dengan benda-benda atau fenomena lain yang telah dikenal, berdasarkan kesamaan atau perbedaan, kedekatan, menurut beberapa standar atau aturan tertentu. Konsep ini biasa disebut dengan Data Mining. Data mining merupakan proses pencarian pola-pola yang menarik dan tersembunyi (hidden pattern) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, data warehouse, atau tempat penyimpanan data lainnya. (Tan dkk, 2006)

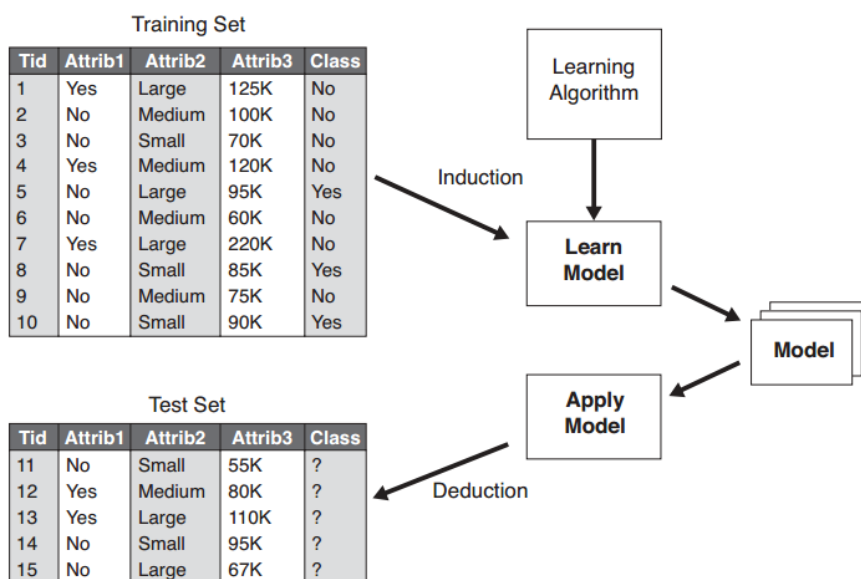
Sistem klasifikasi terbagi menjadi dua, supervised dan unsupervised. Dalam klasifikasi supervised, pemetaan dari satu set input data vektor sampai set kelas diskrit terbatas dimodelkan dalam hal beberapa fungsi matematika, di mana vektor dari parameter disesuaikan. Sedangkan pada klasifikasi unsupervised, disebut clustering atau eksplorasi analisis data, tidak ada data berlabel yang tersedia. Tujuan dari clustering adalah untuk memisahkan data tidak berlabel yang terbatas dan diskrit struktur data tersembunyi (alami). Dengan kata lain Clustering adalah suatu teknik untuk mengklasifikasikan data ke dalam kelompok-kelompok tertentu yang didasarkan atas kesamaan sifat atau karakter. Data yang mempunyai karakteristik yang sama akan dijadikan ke dalam satu kelompok sedangkan data dengan karakteristik berbeda akan masuk atau membentuk kelompok yang lain.

Klasifikasi adalah proses pembelajaran fungsi target f yang memetakan setiap atribut x ke salah satu kelas yang telah ditetapkan label y . Fungsi target biasa disebut dengan model. Definisi klasifikasi dapat dilihat pada Gambar 2.7.



Gambar 2.7 Definisi klasifikasi

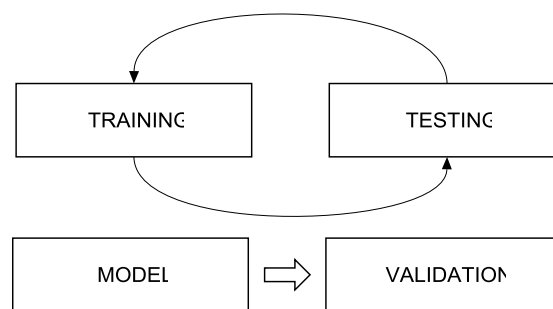
Teknik klasifikasi (classifier) adalah pendekatan sistematis untuk membangun model klasifikasi dari set input data. Misalnya, *Decision Tree* classifiers, rule-based classifiers, neural networks, support vector machines, and naive Bayes classifiers. Setiap teknik menerapkan algoritma pembelajaran untuk mengidentifikasi model terbaik untuk menghubungkan antara atribut input dan kelas output. Tujuan utama dari algoritma pembelajaran adalah untuk membangun model dengan kemampuan generalisasi yang baik; yaitu, model yang akurat memprediksi label kelas dari catatan sebelumnya tidak diketahui. Berikut pendekatan umum untuk memecahkan masalah klasifikasi. Pertama, menyiapkan data set terdiri dari catatan kelas label yang telah dikenal. Training set digunakan untuk membangun model klasifikasi, yang kemudian diterapkan untuk mencari nilai output data test yang kelasnya tidak diketahui. Ilustrasi tersebut dapat dilihat pada Gambar 2.8.



Gambar 2. 8 Pendekatan umum proses klasifikasi.

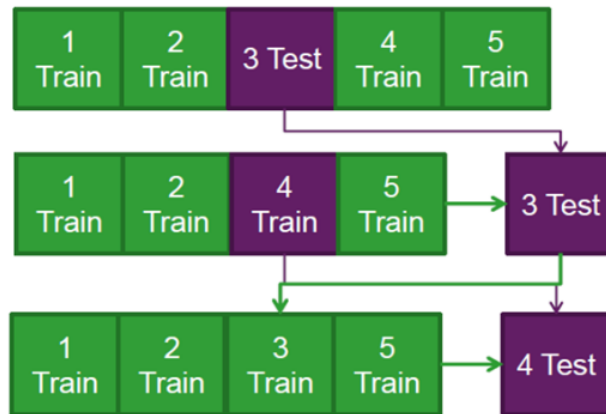
2.2.3.1 Cross-validation

Sebuah alternatif untuk subsampling acak adalah *cross-validation*. Dalam *cross validation* sederhana data dibagi menjadi 2 kelompok secara random, yaitu data training dan data testing. Kemudian dilakukan pertukaran fungsi, yang sebelumnya menjadi data training berubah menjadi data testing, dan sebelumnya data testing berubah menjadi data training. Pendekatan ini disebut dengan *two-fold-cross-validation*. Data training digunakan untuk menemukan model rule dari suatu data set. Model rule tersebut kemudian diuji dengan data testing. Dan dihasilkan nilai akurasi dan error dari model rule tersebut. Total error dihitung dari penjumlahan error kedua proses tersebut. Ilustrasi *two-fold-cross-validation* dapat dilihat pada Gambar 2.9.



Gambar 2.9 *Two-fold-cross-validation*

Cross-validation membagi data secara random. Untuk meminimalkan kesalahan karena proses random maka Cross-validation biasa dilakukan berkali-kali. Cara ini biasa disebut dengan *k-cross validation*. Dengan k adalah konstanta yang membagi jumlah data sebanyak k subsample dan dilakukan uji sebanyak k -kali. Misal $k = 5$, maka data akan dibagi menjadi 5 sama besar. 4 partisi menjadi train dan 1 partisi menjadi test. Setelah uji pertama, maka pada uji kedua data test akan berganti menjadi data training. Pengujian dilakukan sebanyak 5 kali, sesuai dengan nilai k yang digunakan. *k-cross validation* sama dengan *two-fold cross validation* hanya berbeda pada nilai k . Ilustrasi *k-cross validation* dapat dilihat pada Gambar 2.10.



Gambar 2.10 *k-cross validation*

Prosedur *k-cross validation* juga sering disebut dengan leave-one-out, karena mengeluarkan 1 kelompok untuk menjadi data test. Total error diperoleh dengan menjumlahkan semua kesalahan pada k-pengujian. (Tan dkk, 2006)

2.2.3.2 Evaluasi

Evaluasi kinerja model klasifikasi didasarkan pada jumlah catatan benar dan catatan tidak benar yang diprediksi oleh model. jumlah ini ditabulasikan dalam tabel yang dikenal sebagai *confusion matriks*. Tabel 2.2 menggambarkan *confusion matriks* pada klasifikasi biner.

Tabel 2. 2 Confusion Matrix

Confusion Matrix		<i>Predicted Class</i>	
		<i>Class = 1</i>	<i>Class = 0</i>
<i>Actual Class</i>	<i>Class = 1</i>	f_{11}	f_{10}
	<i>Class = 0</i>	f_{01}	f_{00}

Setiap entri f_{ij} dalam tabel ini menunjukkan jumlah record dari kelas i diprediksi menjadi kelas j . Misalnya, f_{01} adalah jumlah record dari kelas 0 salah memperkirakan sebagai kelas 1. Berdasarkan entri dalam confusion matrix, jumlah prediksi yang benar dibuat oleh model ini ($f_{11}+f_{00}$) dan jumlah total prediksi yang salah adalah ($f_{01}+f_{10}$)

Meskipun *confusion matrix* memberikan informasi yang dibutuhkan untuk menentukan seberapa baik model klasifikasi, meringkas informasi ini dengan satu hasil akan membuatnya lebih mudah untuk membandingkan kinerja model yang berbeda. Hal ini dapat dilakukan dengan menggunakan akurasi (*accuracy*), yang didefinisikan sebagai berikut:

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (2.1)$$

Selain dapat dinyatakan dengan akurasi, kinerja model juga dapat dinyatakan dengan tingkat kesalahan (*error rate*).

$$Error Rate = \frac{\text{Number of wrong predictions}}{\text{Total number of predictions}} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (2.2)$$

Semakin kecil tingkat kesalahan maka kinerja model semakin baik. (Tan, Steinbach dan Kumar 2006)

2.2.3.3 Algoritma *Decision Tree*

Salah satu algoritma yang digunakan untuk membangun pohon keputusan yang berbasis algoritma induksi pohon keputusan yaitu C4.5. *Decision Tree* adalah sebuah diagram alir yang berbentuk seperti struktur pohon yang mana setiap internal node menyatakan pengujian terhadap suatu atribut, setiap cabang menyatakan output dari pengujian tersebut dan leaf node menyatakan kelas-kelas atau distribusi kelas. Node yang paling atas disebut sebagai root node atau node akar. Sebuah root node akan memiliki beberapa edge keluar tetapi tidak memiliki edge masuk, internal node akan memiliki satu edge masuk dan beberapa edge keluar, sedangkan leaf node hanya akan memiliki satu edge masuk tanpa memiliki edge keluar.

Algoritma *decision tree* menggunakan entropy dalam proses pemilihan simpul. Entropy menyatakan keragaman data. Entropy digunakan dalam categorical data karena pada data tersebut tidak jarak yang melekat diantara nilai-

nilai data. (Lidkk, 2004). Entropy dihitung dari probabilitas kemungkinan nilai atribut tertentu terhadap kelas tertentu.

Perhitungan entropy menggunakan logaritma. Karena entropy merupakan tipe informasi, dan cara termudah untuk menghitung informasi dijadikan bentuk bit atau byte. Satuan dasar dari informasi adalah bit, yang menyatakan 2 kemungkinan (0 atau 1)

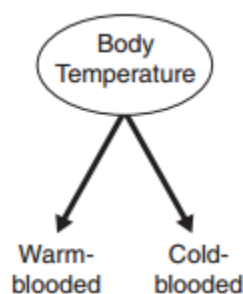
Teori Shannon adalah dasar algoritma ID3 dan begitu juga C4.5. Entropi Shannon adalah yang paling dikenal dan paling banyak digunakan. Teori membandingkan kejadian dengan probabilitas besar dengan kejadian probabilitas kecil atau langka.

Secara umum algoritma C4.5 dalam membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar.
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

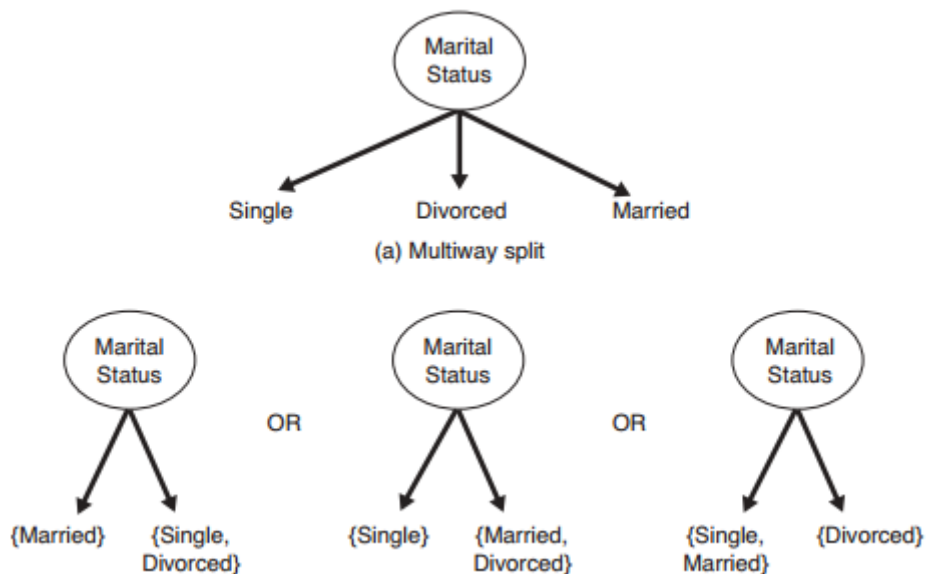
Dalam memilih atribut sebagai akar, ditentukan dari nilai Gain Ratio terbesar dari atribut yang dihitung. Sebelum menghitung nilai gain Ratio dari atribut, hitung dahulu nilai entropy.

Induction pada decision *tree* dilakukan sesuai dengan jenis atribut. Atribut bisa binary attribute, ordinal attribute, nominal attribute atau continuous attribute. pada binary attribute hanya terdapat 2 label sehingga *tree* hanya dibagi menjadi 2. *Induction tree* berdasarkan *binary attribute* ditunjukkan Gambar 2.11.



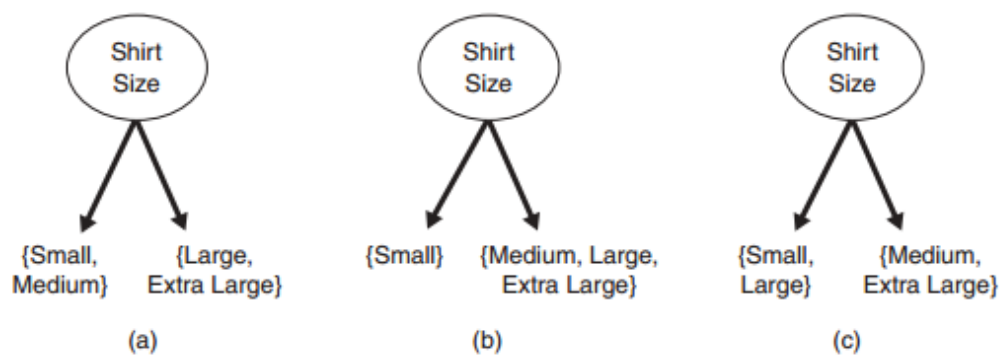
Gambar 2. 11 *Binary Attribute*

Pada *nominal attribute*, jumlah label bisa berjumlah banyak. Ada 2 pendekatan untuk *nominal attribute*, pertama dengan *multiway split*, kedua dengan *binary split*. *Induction tree* berdasarkan nominal attribute ditunjukkan Gambar 2.12.



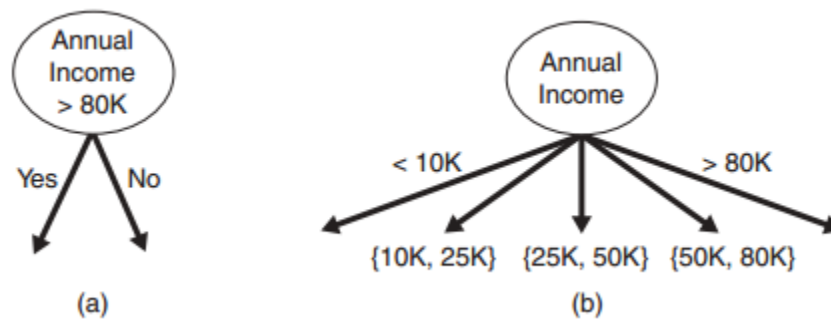
Gambar 2. 12 *Nominal Attribute*

Pada *ordinal attribute*, jumlah label bisa berjumlah banyak. Sama dengan *nominal attribute*, *ordinal attribute* bisa diselesaikan dengan *multiway split* atau *binary split*. *Induction tree* berdasarkan *ordinal attribute* ditunjukkan pada Gambar 2.13.



Gambar 2. 13 *Ordinal Attribute*

Pada continuous attribute, penyelesaian *induction tree* dapat dilakukan dengan *binary split* maupun *multiway split*. Jika hasil kondisi tes terbagi menjadi 2, $A < v$ atau $A > v$, maka menggunakan *binary split*. Jika hasil kondisi tes terbagi menjadi beberapa bagian, maka menggunakan *multiway split*. *Induction tree* berdasarkan *continuous attribute* ditunjukkan Gambar 2.14.



Gambar 2.14 Continuous Attribute

Shannon Entropy

Secara umum, diberikan distribusi probabilitas $P = (p_1, p_2, \dots, p_n)$ pada sampel S , dan informasi yang dibawa oleh distribusi ini sebut juga entropi P . Untuk menghitung nilai entropi digunakan rumus (HSSIN dkk, 2009) :

$$Entropy(S) = - \sum_{i=1}^n p_i (\log_2 p_i) \quad (2.3)$$

Dengan :

- S = Himpunan Kasus
- n = Jumlah kelas partisi S
- p_i = Proporsi dari S_i terhadap S

Secara statistik, entropi menyatakan ukuran ketidakpastian secara probabilistik. Semakin mendekati 0, maka semakin murni (pasti). Sebaliknya, semakin besar nilai entropi maka semakin tidak murni (tidak pasti). Dan setelah menghitung entropi, maka selanjutnya menghitung nilai Gain atau Information Gain.

Information Gain merupakan suatu ukuran korelasi pada model parametrik yang menggambarkan ketergantungan antara dua peubah acak X dan Y. Information Gain memiliki rumus

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.4)$$

dengan :

- S = Himpunan kasus
- A = Atribut
- n = Jumlah partisi atribut A
- $|S_i|$ = Jumlah kasus pada partisi ke i
- $|S|$ = Jumlah kasus dalam S

Perbedaan mendasar antara ID3 dengan C4.5 adalah pada saat menentukan atribut sebagai simpul. Apabila ID3 menggunakan Gain, maka pada C4.5 menggunakan Gain Ratio, perhitungan Gain Ratio didefinisikan sebagai berikut:

$$Gain\ Ratio(S, A) = \frac{Gain(S, A)}{Split\ Information(S, A)} \quad (2.5)$$

Dengan Split information sebagai berikut :

$$Split\ Information(S, A) = \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (2.6)$$

2.2.3.4 Pruning

Setelah membangun *tree*, langkah *pruning* (pemangkasan) dapat dilakukan untuk mengurangi ukuran *tree*. *Tree* yang terlalu besar rentan terhadap fenomena yang dikenal sebagai overfitting. Pemangkasan membantu dengan memotong cabang-cabang pohon awal dengan cara yang meningkatkan kemampuan generalisasi dari pohon keputusan. Untuk mengetahui overfitting atau underfitting dapat dengan mudah dilihat pada proses cross-validation. Jika error validasi dan

error pelatihan keduanya tinggi maka kondisi underfitting. Jika error validasi tinggi sementara error training rendah maka kondisi overfitting. Jika kondisi error validation dan error training sama rendah maka disebut kondisi good fit.

$$e = \frac{f + \frac{z^2}{2N} + z \sqrt{\frac{f}{N} - \frac{f^2}{N} - \frac{z^2}{4N^2}}}{1 + \frac{z^2}{N}} \quad (2.7)$$

dengan:

e = error estimate untuk node

f = error on training data

N = jumlah sampel dalam node

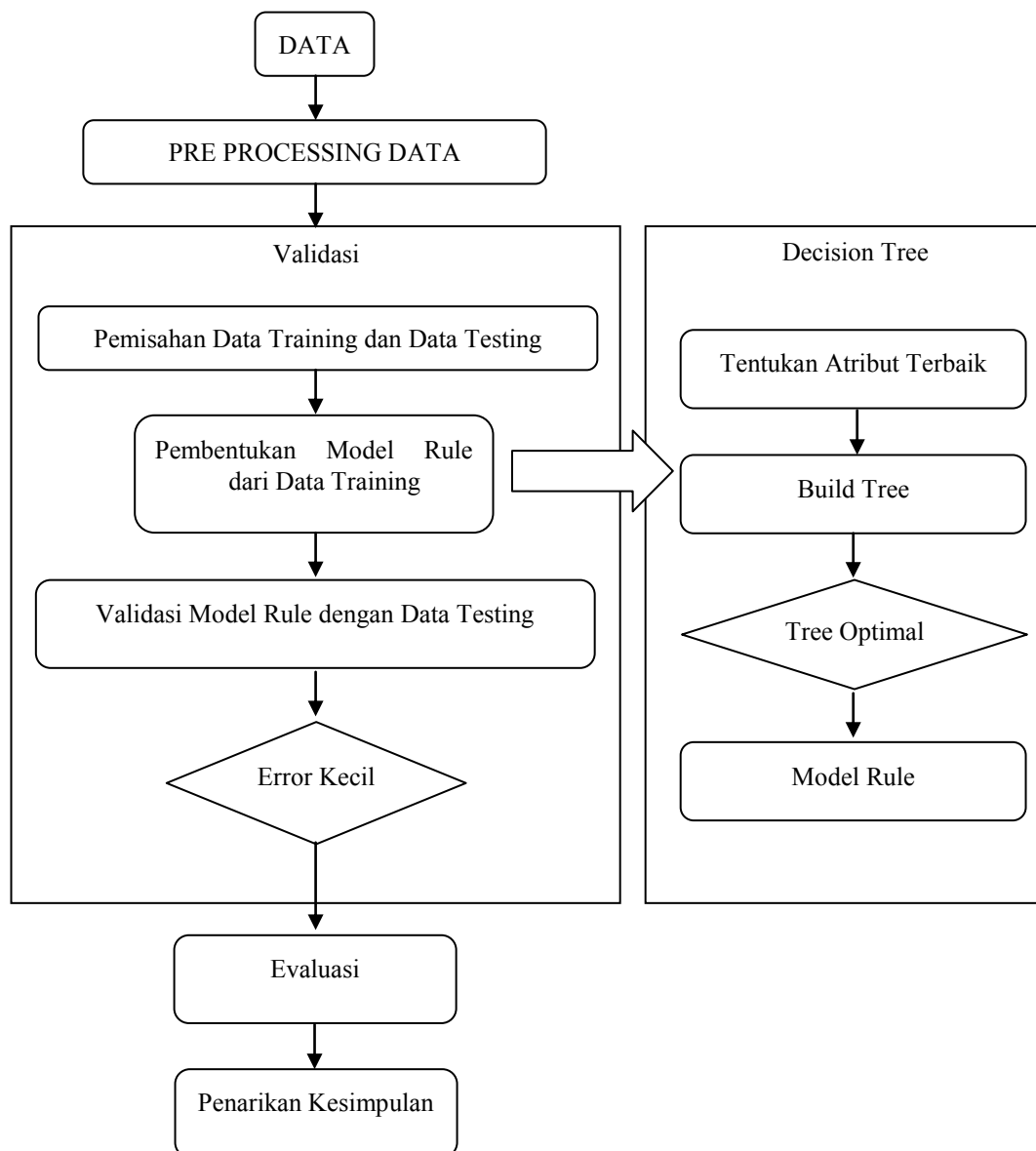
z = upper confidence limit (determined by the confidence interval)

Halaman ini sengaja dikosongkan

BAB III

METODELOGI PENELITIAN

Metodologi Penelitian ini terbagi menjadi 2 tahap yaitu pertama, tahap pre-processing data dan kedua, tahap simulasi decision *tree*. Data diubah menjadi bentuk .csv. Kemudian dilakukan simulasi decision *tree* dengan teknik validasi cross validation dengan aplikasi bantu.



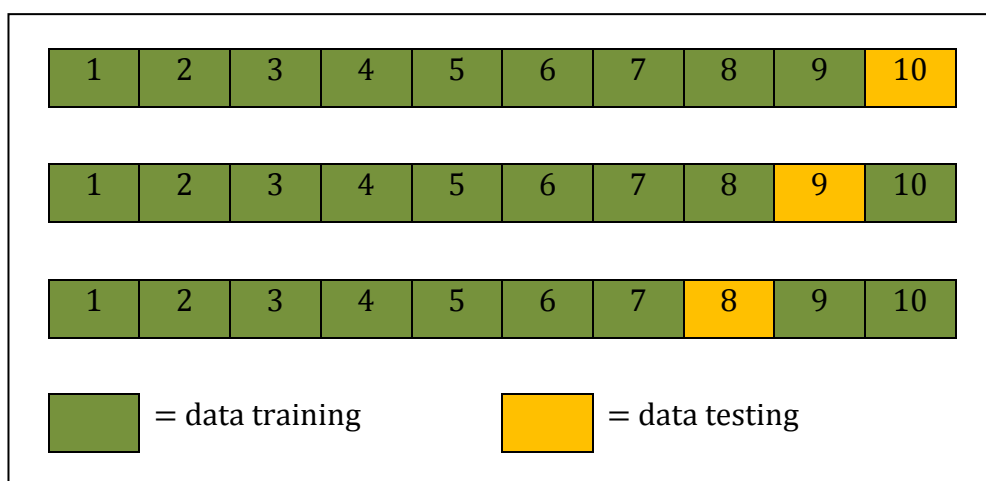
Gambar 3. 1 Metodologi Penelitian

3.1 Tahap Pre-processing Data

Data set dalam penelitian ini diambil dari Dinas Komunikasi dan Informatika Kota Makasar. Data diperoleh tahun 2015. Terdiri dari 10 atribut yang berbeda, dengan 9 atribut berupa input dan 1 atribut berupa target atau output. Data kemudian diolah sehingga data tersebut hanya terdiri dari 1 baris header, 108 baris data dan 10 kolom atribut. Data kemudian disimpan dalam format .csv. Pre-processing data ini dapat menggunakan aplikasi Ms. Excel.

3.2 Validasi

Validasi pada penelitian ini menggunakan metode cross validation. Cross validation yang dilakukan berulang kali biasa disebut dengan k-fold validation. Nilai k menunjukkan jumlah simulasi yang dilakukan. Maka akan dihasilkan pula k model rule. Nilai k yang digunakan pada penelitian ini 10-fold validation. Artinya data akan dibagi menjadi 10 kelompok sama besar atau mendekati sama besar. 1 kelompok akan menjadi data testing, dan 9 kelompok lainnya akan menjadi data training. Data training akan digunakan untuk mencari model rule dengan menggunakan algoritma decision *tree*. Model rule yang terbentuk akan diuji validasi menggunakan data testing. Untuk mempermudah dalam memahami teori cross validation lihat ilustrasi 10 – fold validation pada Gambar 3.2.

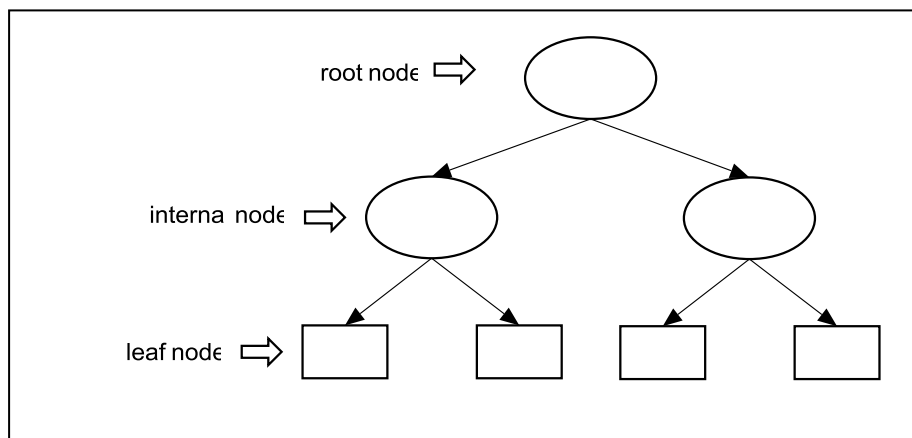


Gambar 3. 2 Ilustrasi 10-fold validation

Pada awal simulasi kelompok ke-10 menjadi data testing, menghasilkan model rule pertama. Setelah itu pada simulasi ke-2, data testing berubah menjadi kelompok ke-9, dan kelompok data yang lain menjadi data training dan dihasilkan model rule ke-2. Dan simulasi dilakukan sampai k kali. Jika nilai k adalah 10, maka simulasi dilakukan sampai 10 kali dan menghasilkan 10 model rule. Tiap model rule yang terbentuk akan divalidasi dengan data testing, dan akan dihitung nilai errornya. Nilai error terkecil yang akan diambil sebagai model rule terbaik.

3.3 Decision tree

Proses penyelesaian klasifikasi dengan metode *decision tree* dimulai dari menentukan root node (simpul akar), internal node atau node percabangan yang memiliki 1 input dan minimal memiliki 2 output, serta terakhir leaf node atau node akhir yang merupakan target klasifikasi. Ilustrasi dari root node, internal node, dan node akhir ditunjukkan pada Gambar 3.3.



Gambar 3. 3 *Induction Tree*

Untuk memudahkan dalam memahami algoritma *decision tree*, pada subbab ini akan dibahas tentang pembentukan keputusan bermain Tenis atau tidak. Keputusan ini akan ditentukan dari 4 atribut input, yaitu cuaca, temperatur, kelembaban, dan angin. Dan target output menjadi 2 kelas, yaitu ya dan tidak. Data Keputusan Permainan Tenis ini berjumlah 14 baris, untuk lebih lengkapnya dapat dilihat pada Tabel 3.1.

Tabel 3. 1 Data Keputusan Permainan Tennis

Cuaca	Temperatur	Kelembaban	Angin	Main atau Tidak
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak
Mendung	Panas	Tinggi	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya
Hujan	Dingin	Normal	Besar	Tidak
Mendung	Dingin	Normal	Besar	Ya
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Hujan	Sedang	Normal	Kecil	Ya
Cerah	Sedang	Normal	Besar	Ya
Mendung	Sedang	Tinggi	Besar	Ya
Mendung	Panas	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Besar	Tidak

Untuk menentukan atribut yang menjadi simpul awal (root node), maka dihitung nilai Gain Ratio masing-masing atribut. Gain Ratio diperoleh dari hasil bagi Information Gain dengan Split Information sesuai persamaan (2.4). berikut urutan dari penyelesaian dengan decision *tree*:

1. Mencari nilai Entropi target atau output
2. Mencari nilai Entropi atribut sesuai label
3. Mencari nilai information Gain masing-masing atribut
4. Mencari nilai Split Information masing-masing atribut
5. Mencari nilai Gain Ratio masing-masing atribut
6. Memilih atribut dengan Gain Ratio terbesar sebagai simpul
7. Kembali ke nomer 1, tanpa menyertakan atribut yang telah menjadi simpul
8. Apabila atribut sudah habis digunakan sebagai node atau Gain Ratio sudah tidak dapat dihitung lagi maka berhenti

Tahap 1

Pada Tabel 3.1 terdapat 14 keputusan, 9 keputusan main dan 5 keputusan tidak main. Entropi output tabel tersebut dapat dihitung dengan persamaan (2.1). perhitungannya sebagai berikut:

$$Entropi = -\frac{9}{14} \log_2 \left(\frac{9}{14} \right) - \frac{5}{14} \log_2 \left(\frac{5}{14} \right) = 0,940 \text{ bits}$$

Setelah menghitung nilai entropi output, maka selanjutnya menghitung nilai entropi masing-masing label pada atribut. Pada atribut cuaca terdapat 3 label, yaitu cerah, mendung, dan hujan. Jumlah label cerah yang beroutput main sejumlah 2, dan label cerah beroutput tidak berjumlah 3. Jumlah label mendung yang beroutput main berjumlah 1, dan label mendung yang beroutput tidak berjumlah 0. Serta label hujan yang beroutput main berjumlah 3, dan label hujan yang beroutput tidak berjumlah 2. Berikut perhitungan entropi dari 3 label pada atribut cuaca.

Perhitungan entropi cuaca = cerah

$$Entropi = -\frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{3}{5} \log_2 \left(\frac{3}{5} \right) = 0,971 \text{ bits}$$

Perhitungan entropi cuaca = mendung

$$Entropi = -\frac{4}{4} \log_2 \left(\frac{4}{4} \right) - 0 \log_2 (0) = 0 \text{ bits}$$

Perhitungan entropi cuaca = hujan

$$Entropi = -\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) = 0,971 \text{ bits}$$

Dari perhitungan diatas didapatkan nilai entropi atribut cuaca untuk masing-masing label. Untuk memudahkan pemahaman, maka perhitungan diatas dimasukkan ke dalam Tabel 3.2.

Tabel 3. 2 Hasil Perhitungan Entropi Cuaca Tahap 1

Atribut	Label	Output		Entropi
		Ya	Tidak	
Cuaca	Cerah	2	3	0,971
	Mendung	4	0	0
	Hujan	3	2	0,971

Setelah menghitung nilai entropi atribut tiap-tiap label, maka selanjutnya menghitung nilai Information Gain dari atribut cuaca dengan persamaan (2.1). Nilai Information Gain dihasilkan dari pengurangan nilai entropi output dikurangi dengan proporsi masing-masing label dalam atribut. Pada Tabel 3.2 , label cerah berjumlah 5, label mendung berjumlah 4, dan label hujan berjumlah 5. Kemudian jumlah label pada seluruh atribut cuaca sebanyak 14. Perhitungan Information Gain pada atribut cuaca adalah sebagai berikut

$$Information\ Gain = 0,940 - \frac{5}{14} (0,971) - \frac{4}{14} (0) - \frac{5}{14} (0,971) = 0,247\ bits$$

Nilai 0,940 merupakan nilai dari entropi output, 5/14 merupakan nilai proporsi dari label cerah. Setelah menghitung nilai Information Gain maka selanjutnya menghitung nilai Split Information. Split information merupakan perhitungan logaritma basis 2 dari jumlah label pada masing-masing atribut.

$$Split\ Information = -\frac{5}{14} \log_2 \left(\frac{5}{14} \right) - \frac{4}{14} \log_2 \left(\frac{4}{14} \right) - \frac{5}{14} \log_2 \left(\frac{5}{14} \right) = 1,577\ bits$$

Setelah mendapatkan nilai split information maka selanjutnya menghitung nilai Gain Ratio. Gain Ratio merupakan hasil dari Gain Information dibagi dengan Split Information.

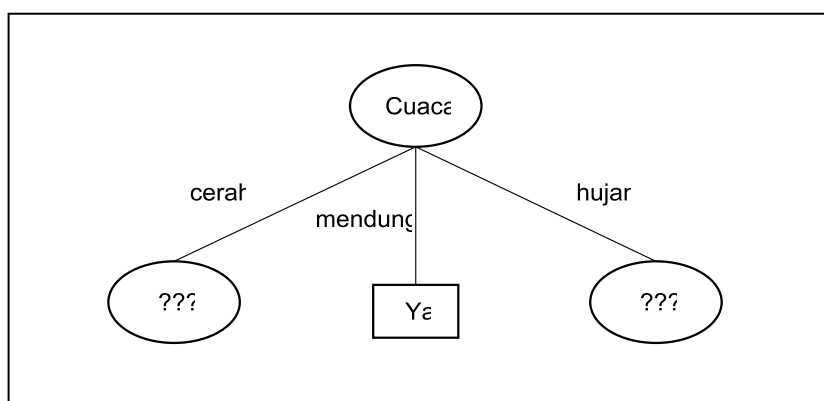
$$Gain\ Ratio = \frac{0,247}{1,577} = 0,156$$

Dari perhitungan diatas didapatkan nilai Information Gain, Split Information, dan Gain Ratio dari atribut cuaca. Cara yang sama digunakan untuk mencari nilai Gain Ratio dari atribut temperatur, kelembaban dan angin. Hasil dari perhitungan pada atribut cuaca, temperatur, kelembaban, dan angin dapat dilihat pada Tabel 3.3.

Tabel 3. 3 Hasil Perrhitungan Gain Ratio Tahap 1

Atribut	Label	Output		Entropi	Information Gain	Split Infomation	Gain Ratio
		Ya	Tidak				
Cuaca	Cerah	2	3	0,971	0,247	1,577	0,156
	Mendung	4	0	0			
	Hujan	3	2	0,971			
Temperatur	Panas	2	3	0,971	0,029	1,557	0,019
	Sedang	4	0	0			
	Dingin	3	2	0,971			
Kelembaban	Tinggi	3	4	0,985	0,152	1	0,152
	Normal	6	1	0,592			
Angin	Besar	6	2	0,811	0,048	0,985	0,049
	Kecil	3	3	1			

Dari Tabel 3.3 terlihat nilai Gain Ratio terbesar ada pada atribut cuaca, selanjutnya atribut cuaca menjadi simpul awal dari *tree* yang akan dibuat.



Gambar 3. 4 Tree Hasil Tahap 1

Tahap 2

Setelah atribut cuaca menjadi simpul awal, maka atribut cuaca tidak lagi dicari nilai Gain Ratio. Pada atribut cuaca terdapat 3 label. Pada label yang nilai entropi 0 maka langsung dibuat keputusan. Sedangkan pada label yang tidak bernilai 0, maka akan dihitung nilai Gain Ratio atribut yang tersisa berdasarkan entropi output dari cabang label. Misal pada cabang label cerah, maka nilai entropi output 0,971 akan dijadikan dasar perhitungan Information Gain. Pada tahap 2, maka data keputusan permainan Tenis berubah, jumlah baris berkurang menjadi 5. Nilai 5 didapatkan dari menyeleksi baris yang hanya berlabel cerah pada kolom cuaca. Untuk lebih jelasnya dapat dilihat pada Tabel 3.4.

Tabel 3. 4 Data Keputusan Permainan Tenis Tahap 2

Cuaca	Temperatur	Kelembaban	Angin	Main atau Tidak
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Cerah	Sedang	Normal	Besar	Ya

Dari tabel 3.4 kemudian dicari nilai gain ratio pada masing-masing atribut seperti pada tahap 1 (selain atribut cuaca).

Perhitungan entropi kelembaban = tinggi

$$Entropi = -\frac{0}{3}\log_2\left(\frac{0}{3}\right) - \frac{3}{3}\log_2\left(\frac{3}{3}\right) = 0 \text{ bits}$$

Perhitungan entropi kelembaban = normal

$$Entropi = -\frac{2}{2}\log_2\left(\frac{2}{2}\right) - \frac{0}{2}\log_2\left(\frac{0}{2}\right) = 0 \text{ bits}$$

Dari perhitungan diatas didapatkan nilai entropi atribut kelembaban untuk masing-masing label. Untuk memudahkan pemahaman, maka perhitungan diatas dimasukkan ke dalam Tabel 3.5.

Tabel 3. 5 Hasil Perhitungan Entropi Kelembaban Tahap 2

Atribut	Label	Output		Entropi
		Ya	Tidak	
Kelembaban	Tinggi	0	3	0
	Normal	2	0	0

Setelah menghitung nilai entropi atribut tiap-tiap label, maka selanjutnya menghitung nilai Information Gain dari atribut kelembaban dengan persamaan (2.1). Pada Tabel 3.5 , label tinggi berjumlah 3 dan label normal berjumlah 2. Kemudian jumlah label pada seluruh atribut kelembaban sebanyak 5. Perhitungan Information Gain pada atribut kelembaban adalah sebagai berikut

$$Information\ Gain = 0,971 - \frac{3}{5} (0) - \frac{2}{5} (0) = 0,971\ bits$$

Nilai 0,971 merupakan nilai dari entropi output, $\frac{3}{5}$ merupakan nilai proporsi dari label tinggi. Setelah menghitung nilai Information Gain maka selanjutnya menghitung nilai Split Information. Split information merupakan perhitungan logaritma basis 2 dari jumlah label pada masing-masing atribut

$$Split\ Information = -\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) = 0,971\ bits$$

Setelah mendapatkan nilai split information maka selanjutnya menghitung nilai Gain Ratio. Gain Ratio merupakan hasil dari Gain Information dibagi dengan Split Information.

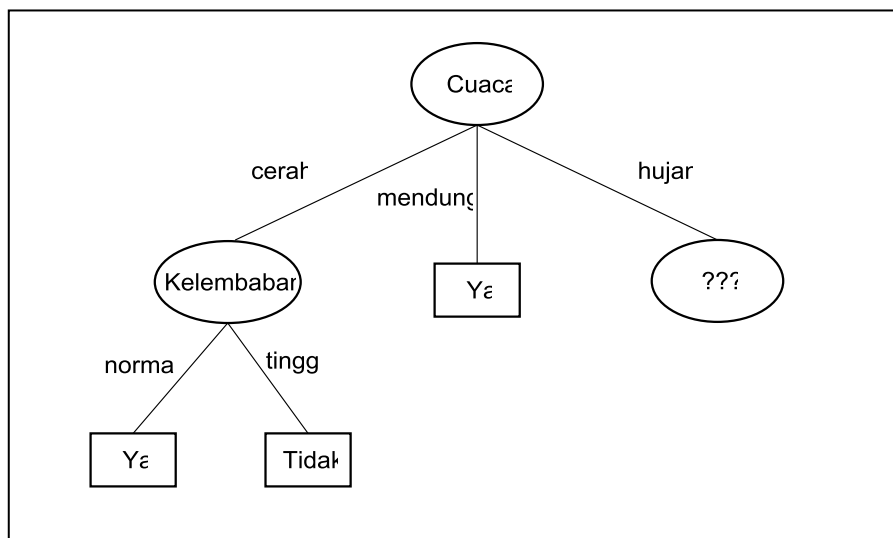
$$Gain\ Ratio = \frac{0,971}{0,971} = 1\ bits$$

Selanjutnya dengan cara yang sama dicari nilai Entropi, Information Gain, Split Information, dan Gain Ratio untuk atribut temperatur dan angin. Hasil perhitungan Entropi, Information Gain, Split Information, dan Gain Ratio atribut kelembaban, temperatur dan angin dapat dilihat pada Tabel 3.6.

Tabel 3. 6 Hasil Perhitungan Gain Ratio Tahap 2

Atribut	Label	Output		Entropi	Information Gain	Split Infomation	Gain Ratio
		Ya	Tidak				
Kelembaban	Tinggi	0	3	0	0,971	0,971	1
	Normal	2	0	0			
Temperatur	Panas	0	2	0	0,571	1,522	0,375
	Sedang	1	1	1			
	Dingin	1	0	0			
Angin	Besar	1	2	0,918	0,020	0,971	0,021
	Kecil	1	1	1			

Dari Tabel 3.6 dapat diketahui bahwa atribut kelembaban memiliki Gain ratio terbesar. Maka pada cabang atribut cuaca label cerah dipilihlah atribut kelembaban. *Tree* tahap 2 ditunjukkan pada Gambar 3.5.



Gambar 3. 5 Tree Hasil Tahap 2

Entropi pada atribut kelembaban bernilai 0 untuk tiap labelnya, maka leaf atau daun yang didapat merupakan daun yang murni sehingga hasilnya bersifat final.

Tahap 3

Setelah atribut kelembaban menjadi simpul tengah (internal node) pada cabang/label cerah, maka selanjutnya dicari simpul tengah (internal node) pada cabang/label hujan. Cara perhitungan tahap 3 sama dengan tahap 2. Pada tahap 3, maka data keputusan permainan Tennis berubah. Data cuaca yang digunakan hanya yang hujan. Lebih jelasnya dapat dilihat pada Tabel 3.7

Tabel 3. 7 Data Keputusan Permainan Tennis Tahap 3

Cuaca	Temperatur	Kelembaban	Angin	Main atau Tidak
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya
Hujan	Dingin	Normal	Besar	Tidak
Hujan	Sedang	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Besar	Tidak

Dari tabel 3.4 kemudian dicari nilai gain ratio pada masing-masing atribut seperti pada tahap 2 (selain atribut cuaca).

Perhitungan entropi angin = panas

$$Entropi = -\frac{2}{5} \log_2 \left(\frac{2}{5} \right) - \frac{3}{5} \log_2 \left(\frac{3}{5} \right) = 0,971 \text{ bits}$$

Perhitungan entropi angin = sedang

$$Entropi = -\frac{4}{4} \log_2 \left(\frac{4}{4} \right) - 0 \log_2 (0) = 0 \text{ bits}$$

Dari perhitungan diatas didapatkan nilai entropi atribut angin untuk masing-masing label. Untuk memudahkan pemahaman, maka perhitungan diatas dimasukkan ke dalam Tabel 3.8.

Tabel 3. 8 Hasil Perhitungan Entropi Kelembaban Tahap 3

Atribut	Label	Output		Entropi
		Ya	Tidak	
Angin	Besar	0	3	0
	Kecil	2	0	0

Setelah menghitung nilai entropi atribut tiap-tiap label, maka selanjutnya menghitung nilai Information Gain dari atribut angin dengan persamaan (2.1). Nilai Information Gain dihasilkan dari pengurangan nilai entropi output dikurangi dengan proporsi masing-masing label dalam atribut. Pada Tabel 3.8 , label besar berjumlah 3 dan label kecil berjumlah 2. Kemudian jumlah label pada seluruh atribut angin sebanyak 5. Perhitungan Information Gain pada atribut angin adalah sebagai berikut

$$Information\ Gain = 0,971 - \frac{3}{5} (0) - \frac{2}{5} (0) = 0,971\ bits$$

Nilai 0,971 merupakan nilai dari entropi output, $\frac{3}{5}$ merupakan nilai proporsi dari label besar. Setelah menghitung nilai Information Gain maka selanjutnya menghitung nilai Split Information. Split information merupakan perhitungan logaritma basis 2 dari jumlah label pada masing-masing atribut

$$Split\ Information = -\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) = 0,971\ bits$$

Setelah mendapatkan nilai split information maka selanjutnya menghitung nilai Gain Ratio. Gain Ratio merupakan hasil dari Gain Information dibagi dengan Split Information.

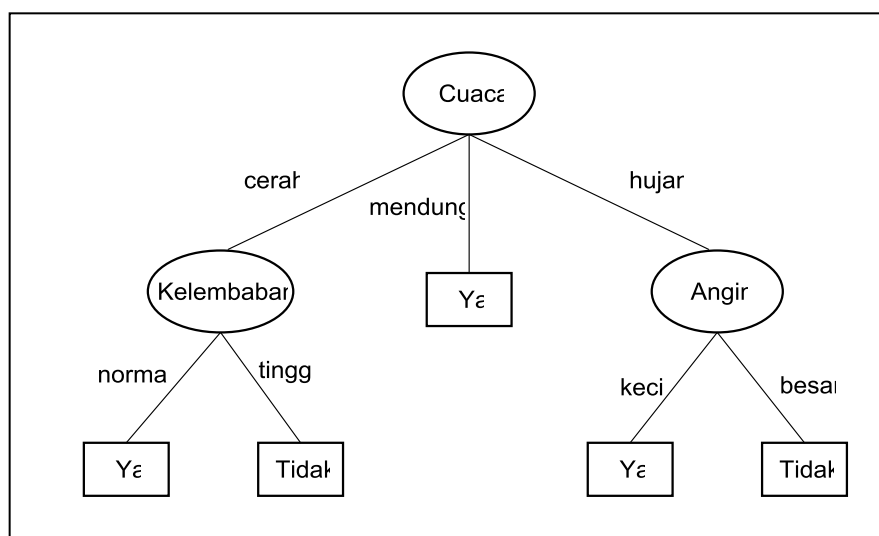
$$Gain\ Ratio = \frac{0,971}{0,971} = 1$$

Selanjutnya dengan cara yang sama dicari nilai Entropi, Information Gain, Split Information, dan Gain Ratio untuk atribut temperatur dan kelembaban. Hasil perhitungan Entropi, Information Gain, Split Information, dan Gain Ratio atribut kelembaban, temperatur dan angin dapat dilihat pada Tabel 3..

Tabel 3. 9 Hasil Perhitungan Gain Ratio Tahap 3

Atribut	Label	Output		Entropi	Information Gain	Split Infomation	Gain Ratio
		Ya	Tidak				
Kelembaban	Tinggi	1	1	1	0,020	0,0971	0,021
	Normal	2	1	0,92			
Temperatur	Panas	0	0	0	0,020	0,442	0,045
	Sedang	2	1	0,92			
	Dingin	1	1	1			
Angin	Besar	0	2	0	0,971	0,971	1
	Kecil	3	0	0			

Dari Tabel 3.9 dapat diketahui bahwa atribut angin memiliki Gain ratio terbesar. Maka pada cabang atribut cuaca label hujan dipilihlah atribut angin. Selanjutnya di buat *tree* tahap 3 seperti ditunjukkan pada Gambar 3.6.

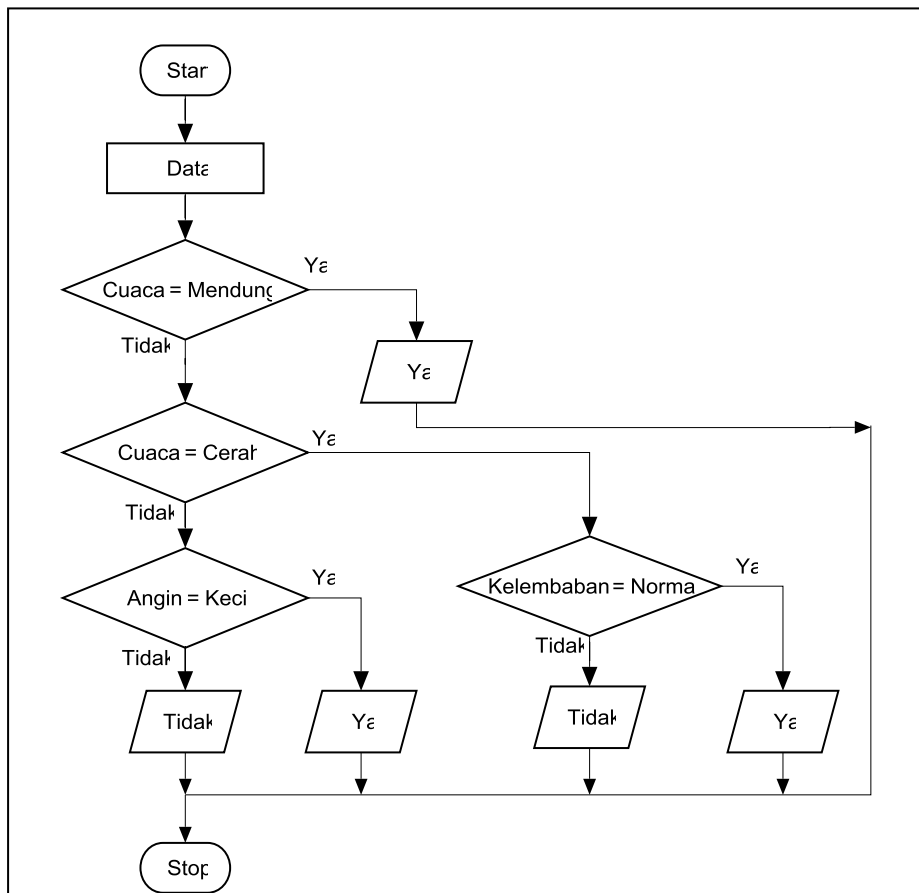


Gambar 3.6 Tree Hasil Tahap 3

Tree tahap 3 telah memiliki leaf akhir, sehingga tidak ada tahap pencarian gain ratio untuk tahap 4. Dari perhitungan ini didapatkan 3 atribut yang digunakan dalam pembentukan *tree*, yaitu cuaca, kelembaban, dan angin. Sedangkan Atribut temperatur tidak berpengaruh dalam pembentukan *tree*.

3.4 Evaluasi

Penentuan ya atau tidak dari permainan tenis dilihat pertama kali pada atribut cuaca, apabila cuaca mendung maka ya bermain tenis. Jika cuaca cerah, maka dilihat atribut kelembaban, apabila kelembaban normal maka ya bermain tenis, bila kelembaban tinggi maka tidak bermain tenis. Dan jika cuaca hujan, maka dilihat atribut angin, apabila angin kecil maka ya bermain tenis, bila angin besar maka tidak bermain tenis. Rule ini dapat dilihat dengan mudah bila dinyatakan dalam diagram alir (flowchart). Flowchart data keputusan bermain tenis dapat dilihat pada Gambar 3.7.



Gambar 3. 7 Flow Chart Data Keputusan Bermain Tenis

Dari model rule diatas kemudian divalidasi dengan data testing menghasilkan confusion matrik sebagai berikut

Tabel 3. 10 Confusion Matrix Data Keputusan Permainan Tennis

Confusion Matrix		<i>Predicted Class</i>	
		<i>Ya</i>	<i>Tidak</i>
<i>Actual Class</i>	<i>Ya</i>	3	0
	<i>Tidak</i>	2	0

Evaluasi suatu klasifikasi dapat dinilai dari akurasi maupun error rate. Persamaan (2.1) merupakan rumus perhitungan accuracy dan persamaan (2.2) merupakan rumus perhitungan error rate. Berikut perhitungan accuracy dan error rate data keputusan permainan tenis

$$Accuracy == \frac{3 + 0}{3 + 0 + 2 + 0} = 0,6 = 60\%$$

$$Error Rate = \frac{0 + 2}{3 + 0 + 2 + 0} = 0,40 = 40\%$$

Dari perhitungan diatas, diketahui nilai accuracy sebesar 60% dan nilai error rate sebesar 40%.

Halaman ini sengaja dikosongkan

BAB IV

ANALISA DAN PEMBAHASAN

4.1 Analisa Data Awal

Data set mengandung identitas kamera dari 108 kamera CCTV. Target atribut yang dicari adalah tipe Kamera CCTV yang terpasang apakah merupakan tipe A, tipe B , atau tipe C. Data terdiri 9 input (independent) atribut. Keseluruhan atribut terlihat pada Tabel 4.1.

Tabel 4. 1 Dataset Kamera CCTV

No.	Atribut	Model Role	Tipe Atribut	Deskripsi/ Label
1	Kamera	Target	Categorical- Nominal	A, B,C
2	Rawan Kriminal	Input	Categorical - Ordinal	Bukan, Jarang, Sedang, Sering
3	Rawan Kecelakaan	Input	Categorical - Ordinal	Bukan, Jarang, Sedang, Sering
4	Rawan Kemacetan	Input	Categorical - Ordinal	Tinggi, Jarang, Sedang, Tinggi
5	Temu Lalu Lintas	Input	Categorical - Nominal	Ya, Tidak
6	Lurus	Input	Categorical - Nominal	Bukan, Alternatif Utama
7	Pertigaan	Input	Categorical - Nominal	Bukan, Alternatif Utama
8	Perempatan	Input	Categorical - Nominal	Bukan, Alternatif Utama
9	Bundaran	Input	Categorical - Nominal	Bukan, Alternatif Utama
10	Lorong	Input	Categorical - Nominal	Bukan, Alternatif Utama

Atribut-atribut tersebut kemudian diuji secara statistik untuk menentukan penyebaran labelnya. Dari uji tersebut diketahui bahwa masing-masing atribut input hanya mempunyai 2 label. Pada algoritma decision *tree* jumlah label menentukan jenis *induction tree*. Jika atribut memiliki 2 label, maka *induction tree* berbentuk *binary tree*. Karena semua atribut input memiliki jumlah label 2, maka *induction tree* pada dataset kamera berbentuk *binary tree*. Label-label yang digunakan pada masing-masing atribut input dapat dilihat pada Tabel 4.2.

Tabel 4. 2 Label Dataset Kamera

No.	Atribut	Label
1	Rawan Kriminal	Sering, Bukan
2	Rawan Kecelakaan	Sering, Bukan
3	Rawan Kemacetan	Tinggi, Bukan
4	Temu Lalu Lintas	Ya, Tidak
5	Lurus	Utama, Bukan
6	Pertigaan	Utama Bukan,
7	Perempatan	Utama,Bukan
8	Bundaran	Utama,Bukan
9	Lorong	Utama, Bukan

4.2 Analisa Performa Tipe CCTV

Dalam penelitian kali ini akan digunakan 3 jenis kamera CCTV, yaitu tipe A, tipe B, dan tipe C. Masing-masing tipe memiliki karakteristik yang berbeda. Misalnya pada kamera tipe A dan tipe C, kedua tipe ini dapat bergerak secara vertikal dan horizontal. Hal ini dapat berguna untuk mengawasi area yang luas. Sedangkan kamera tipe B, hanya dapat bergerak secara vertikal. Spesifikasi lengkap dari kamera CCTV tipe A, tipe B, dan C dapat dilihat pada Tabel 4.2.

Pada baris kedua Tabel 4.2 terlihat perbedaan tipe kamera ditinjau dari kualitas rekaman video yang dihasilkan. Untuk kualitas 4CIF resolusi video sebesar 704×480 piksel. Sedangkan kamera dengan kualitas 1,3 MP dapat

menyimpan video dengan resolusi sebesar 1280×724 piksel. Ukuran piksel yang besar membuat video dapat menangkap gambar lebih detil. Tipe kamera yang digunakan pada penelitian ini semuanya berbasis IP Camera, dengan begitu tidak membutuhkan DVR analog. Keuntungan tidak menggunakan DVR analog adalah hasil rekaman dapat langsung disimpan ke dalam memory penyimpanan tanpa harus dikonversi menjadi digital terlebih dahulu. DVR analog memiliki keterbatasan dalam penyimpanan, yaitu hanya mampu menyimpan file ukuran D1 (720x576 piksel). Berapapun ukuran lensa kamera yang digunakan maka nantinya hanya bisa tersimpan maksimal berukuran 720 x 576 piksel.

Tabel 4. 3 Spesifikasi CCTV Tipe A, Tipe B, dan Tipe C

Tipe A	Tipe B	Tipe C
Pan, tilt, Zoom 360° Endless	Bullet Fixed Camera with IR	Pan, tilt, zoom 360° endless
4CIF, H.264, 30fps	1.3MP H.264, 30fps	1.3MP, H.264, 30fps
Lens 3.69~44.32 mm(12x opt zoom)	Lens 3~8.5mm	Lens 3.5 ~ 150.5mm(43x opt zoom)
Angular field of view H: 54.4° (Wide); 4.62° (tele) V: 42.32° (wide); 3.58° (tele)	Angular field of view H: 93.4° (wide); 33.1° (Tele) V: 73.8° (wide); 26.5° (tele)	Angular field of view H: 53.92° (wide); 1.396° (tele) V: 44.08° (wide); 1.12° (tele)
Day& Night Min illumination Color 0.7lux ; B/W 0.07lux	Day & Night IR On 0 lux	Day& Night Min illumination Color 0.7lux ; B/W 0.07lux

Pada baris ketiga menampilkan perbedaan ukuran lensa dari masing-masing kamera. Ukuran lensa ini juga berpengaruh pada angle of view (sudut pandang) dari kamera. Dengan bukaan lensa yang semakin besar, maka sudut pengambilan gambar dari kamera semakin kecil namun dapat melihat semakin detil untuk objek yang letaknya jauh. Hal ini serupa dengan mata manusia, apabila mata terkonsentrasi untuk melihat benda yang jauh letaknya, maka sudut pandang

mata menjadi semakin kecil (ilustrasi dapat dilihat pada Gambar 2.6). Kamera dengan ukuran lensa dan sudut pandang yang besar dapat digunakan di area yang luas.

Pada baris kelima menampilkan sensitivitas dari kamera. Sensitivitas ini tertulis dalam satuan *lux*. Semakin mendekati 0, maka kamera semakin sensitiv karena dapat menangkap dengan intensitas cahaya yang sangat minim (sangat gelap). Sebaliknya, semakin besar nilai *lux* dari kamera, maka sensitivitas kamera semakin berkurang karena kamera semakin membutuhkan intensitas cahaya

Fungsi PTZ sangat berkaitan dengan real time monitoring. Misal, apabila terjadi kemacetan akibat timer traffic light yang tidak sesuai, maka bisa langsung disesuaikan dengan melihat kondisi sekitar. Begitu juga dengan lensa yang berukuran besar, sehingga dapat melihat dengan jelas obyek yang letaknya jauh tanpa mengurangi ketajaman obyek. Hal ini diperlukan, misalnya bila terjadi tindak kriminalitas maka peetugas pengawas CCTV dapat dengan segera memfokuskan kamera ke nomor kendaraan pelaku kriminalitas, sehingga memudahkan dalam penangkapan pelaku.

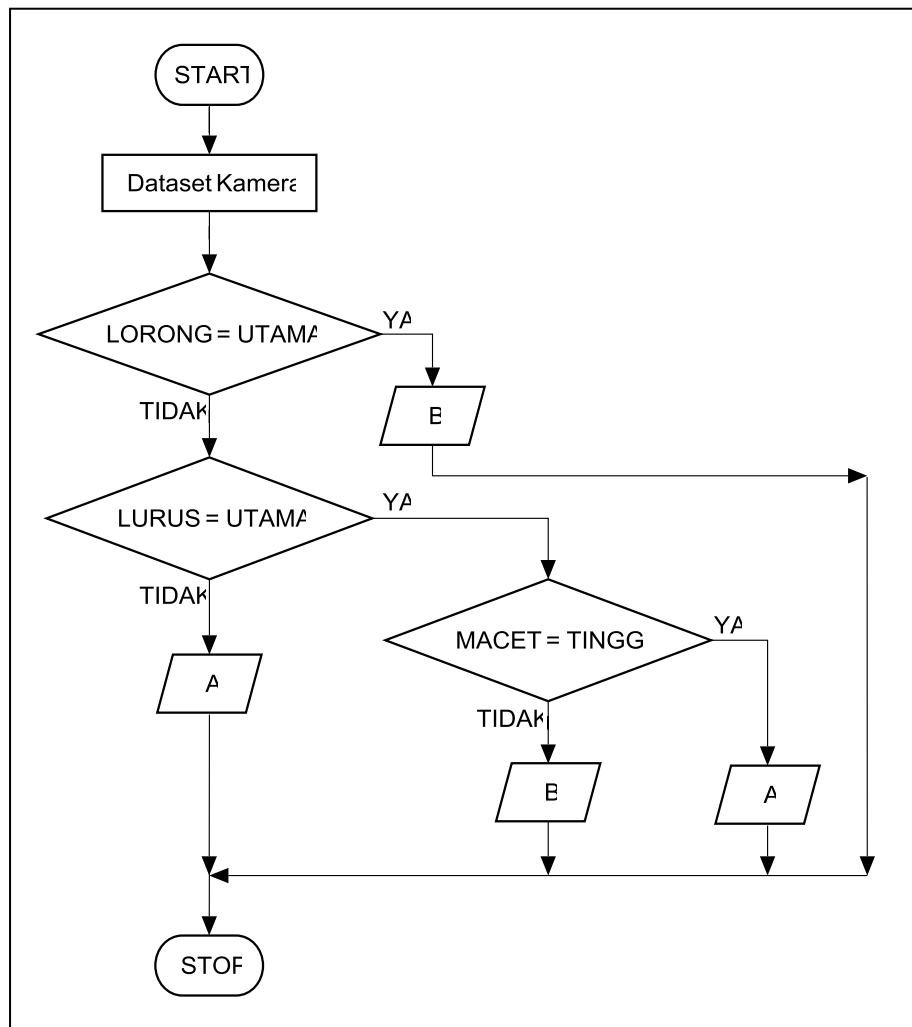
4.3 *Klasifikasi Decision Tree*

Pemilihan decision *tree* pada penelitian kali ini terkait erat dengan jenis dataset yang akan digunakan. Terlihat pada Tabel 4.1 dataset kamera bersifat categorical. Dimana pada categorical data tidak memiliki jarak yang melekat diantara data yang satu dengan yang lainnya. Oleh sebab itu dipilihlah perhitungan berbasis entropi. Entropi merupakan pengukuran terhadap keragaman data. Oleh karena itu dipilihkan algoritma decision *tree* yang prosesnya menggunakan entropi.

Metodologi penelitian ini menggunakan validasi dengan k-cross validation. Sehingga pada saat mensimulasikan data dengan algoritma C4.5 maka akan terdapat k *tree* (model rule). Dalam penelitian ini k=10, sehingga data akan dibagi menjadi 10 kelompok, dengan 1 kelompok sebagai data testing dan 9 kelompok sebagai data training. Pengelompokkan ini dilakukan secara random. Berikut 10 model rule hasil simulasi algoritma C4.5 :

Rule 1

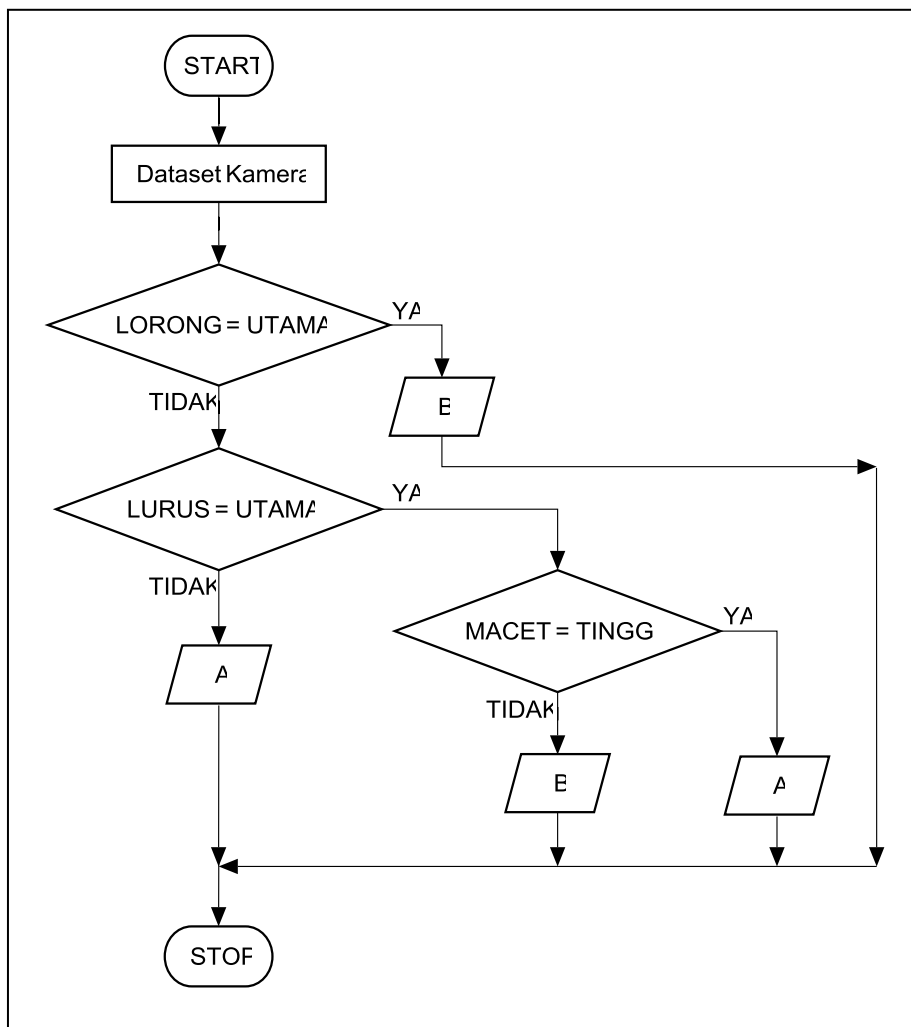
Tree ke-1 hasil simulasi algoritma decision *tree* C4.5 melibatkan 3 atribut dalam pembuatan model rule. Atribut tersebut yaitu lorong, lurus, dan macet. Apabila lorong merupakan lorong utama maka kamera tipe B, bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut macet. Bila macet tinggi maka kamera tipe A, sebaliknya bila bukan jalan macet maka kamera tipe B. Pada rule pertama tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-1, dapat dilihat pada flow chart Gambar 4.1.



Gambar 4. 1 Rule ke-1 Dataset Kamera

Rule 2

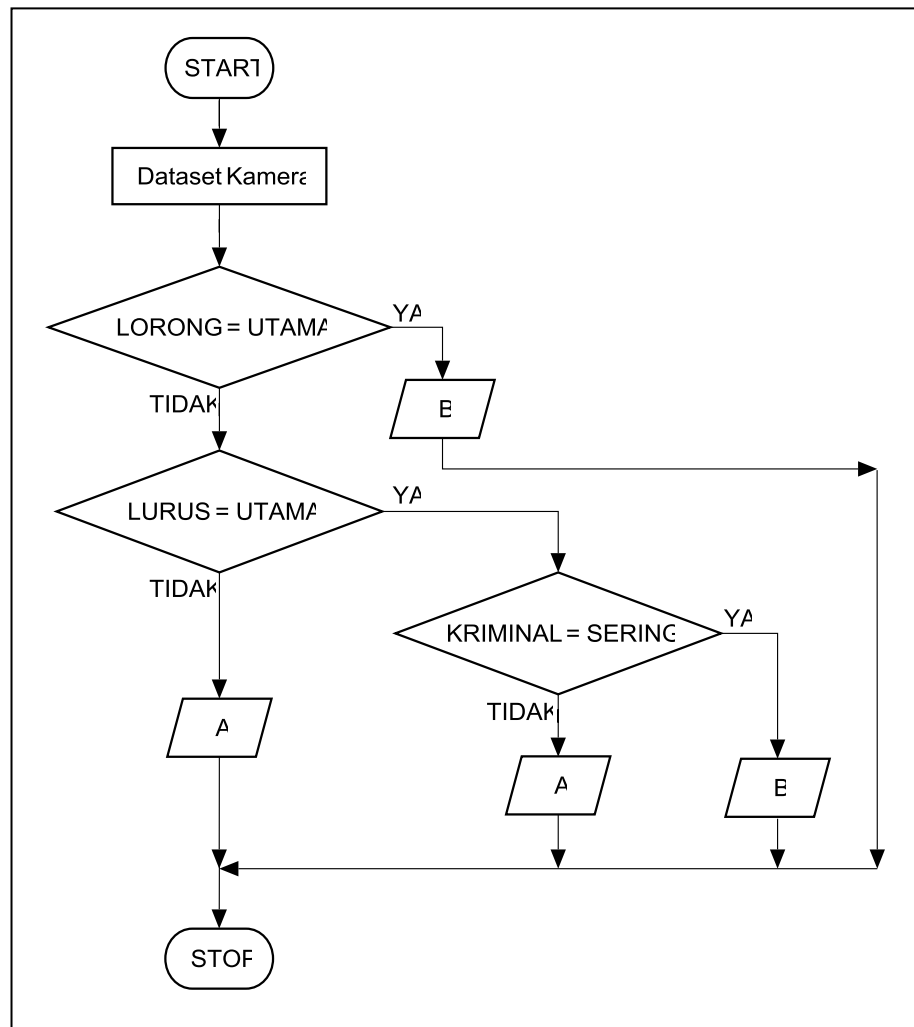
Tree ke-2 hasil simulasi algoritma decision *tree* C4.5 melibatkan 3 atribut dalam pembuatan model rule. Atribut tersebut yaitu lorong, lurus, dan macet. Apabila lorong merupakan lorong utama maka kamera tipe B, bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut macet. Bila macet tinggi maka kamera tipe A, sebaliknya bila bukan jalan macet maka kamera tipe B. Pada rule pertama tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-2, dapat dilihat pada flow chart Gambar 4.2.



Gambar 4. 2 Rule Ke-2 Dataset Kamera

Rule 3

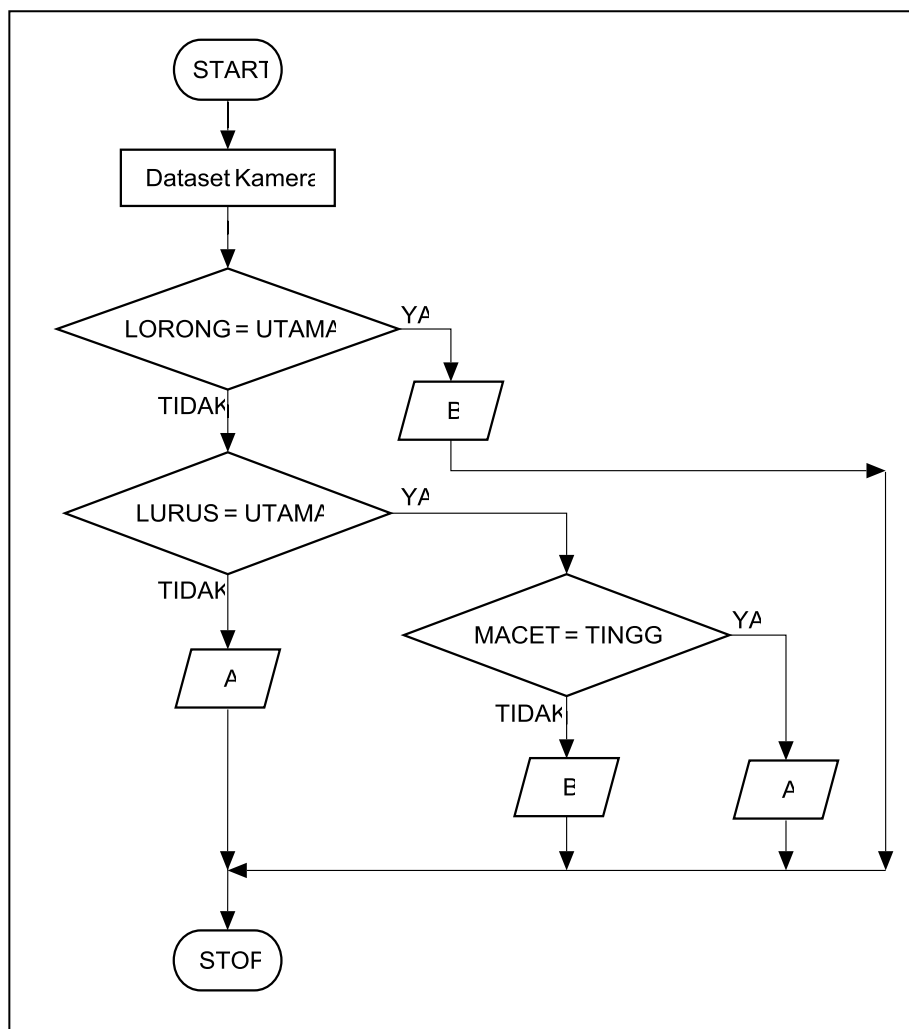
Tree ke-3 hasil simulasi algoritma decision tree C4.5 melibatkan 3 atribut dalam pembentukan model rule. Atribut tersebut yaitu lorong, lurus, dan kriminal. apabila lorong merupakan lorong utama maka kamera tipe B. Bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut kriminal. bila kriminal sering maka kamera tipe B. Sebaliknya bila jalan bukan rawan kriminal maka kamera tipe A. Pada rule ke-3 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-3, dapat dilihat pada flowchart Gambar 4.3.



Gambar 4. 3 Rule Ke-3 Dataset Kamera

Rule 4

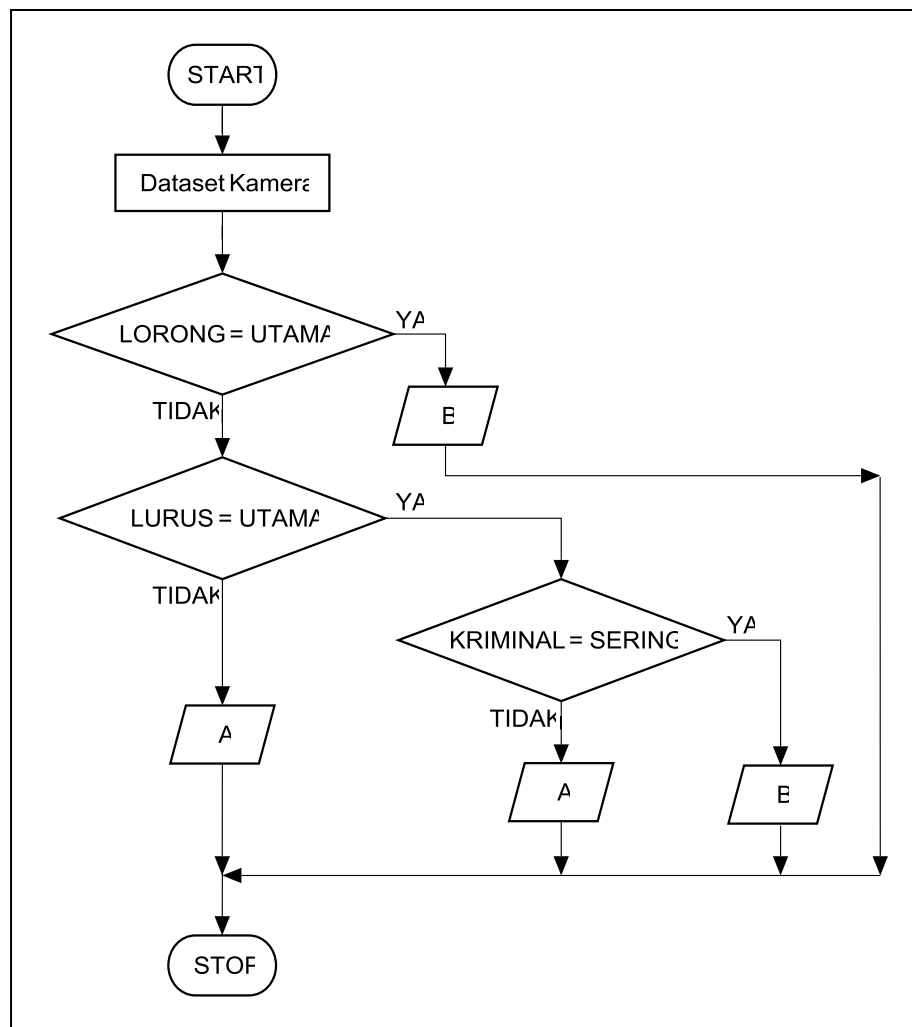
Tree ke-4 hasil simulasi algoritma decision *tree* C4.5 melibatkan 3 atribut dalam pembuatan model rule. Atribut tersebut yaitu lorong, lurus, dan macet. Apabila lorong merupakan lorong utama maka kamera tipe B, bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut macet. Bila macet tinggi maka kamera tipe A, sebaliknya bila bukan jalan macet maka kamera tipe B. Pada rule ke-4 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-4, dapat dilihat pada flow chart Gambar 4.4.



Gambar 4. 4 Rule Ke-4 Dataset Kamera

Rule 5

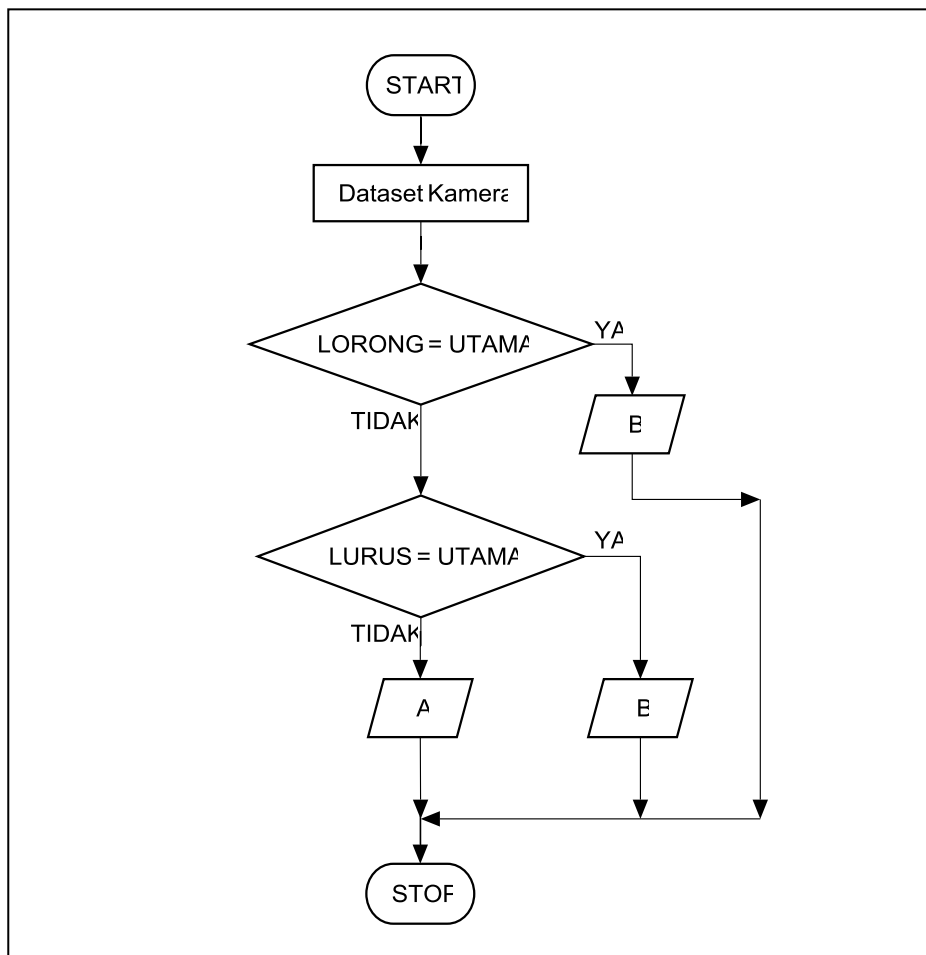
Tree ke5 hasil simulasi algoritma decision tree C4.5 melibatkan 3 atribut dalam pembentukan model rule. Atribut tersebut yaitu lorong, lurus, dan kriminal. Apabila lorong merupakan lorong utama maka kamera tipe B. Bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut kriminal. bila kriminal sering maka kamera tipe B. Sebaliknya bila jalan bukan rawan kriminal maka kamera tipe A. Pada rule ke-5 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-5, dapat dilihat pada flowchart Gambar 4.5.



Gambar 4. 5 Rule Ke-5 Dataset Kamera

Rule 6

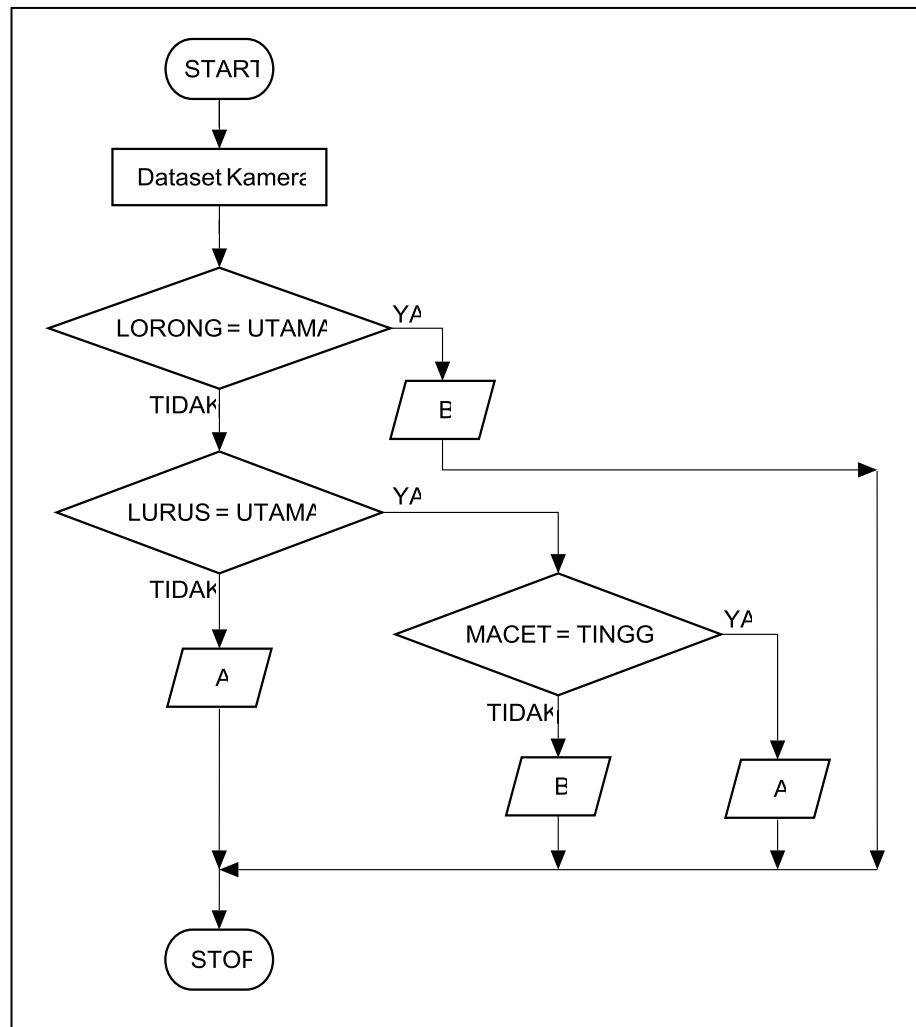
Tree ke-6 hasil simulasi algoritma decision *tree* C4.5 melibatkan 2 atribut dalam pembentukan model rule. Atribut tersebut yaitu lorong dan lurus. Apabila lorong merupakan lorong utama maka kamera yang digunakan adalah kamera tipe B. Bila lorong bukan lorong utama maka dipersyaratkan atribut selanjutnya, yaitu atribut lurus. Bila lurus merupakan utama maka kamera yang digunakan adalah kamera tipe B. Sebaliknya bila bukan merupakan jalan lurus maka menggunakan kamera tipe A. Ada perbedaan yang mencolok dari rule ke-6, yaitu hanya melibatkan 2 atribut. sedangkan rule yang lain melibatkan 3 atribut. Pada rule ke-6 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-6, dapat dilihat pada flow chart Gambar 4.6.



Gambar 4. 6 Rule Ke-6 Dataset Kamera

Rule 7

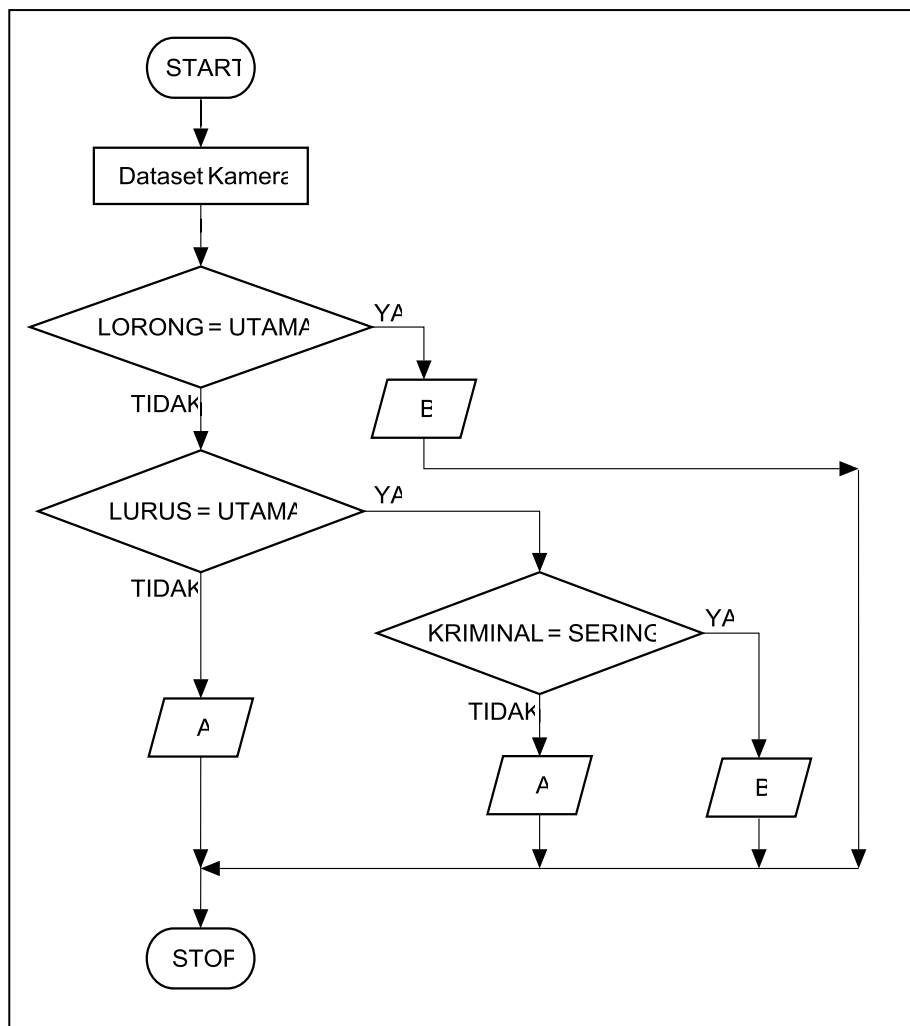
Tree ke-7 hasil simulasi algoritma decision *tree* C4.5 melibatkan 3 atribut dalam pembuatan model rule. Atribut tersebut yaitu lorong, lurus, dan macet. Apabila lorong merupakan lorong utama maka kamera tipe B, bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut macet. Bila macet tinggi maka kamera tipe A, sebaliknya bila bukan jalan macet maka kamera tipe B. Pada rule ke-7 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-7, dapat dilihat pada flow chart Gambar 4.7.



Gambar 4. 7 Rule Ke-7 Dataset Kamera

Rule 8

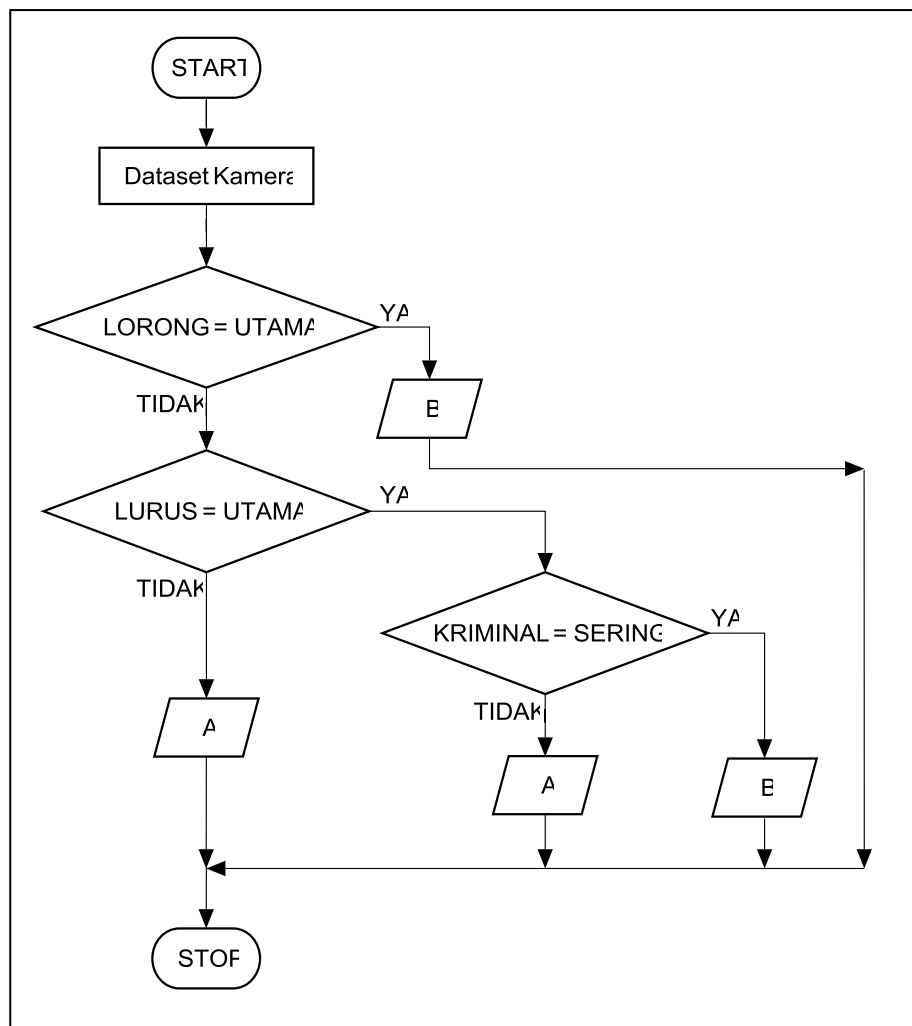
Tree ke-8 hasil simulasi algoritma decision *tree* C4.5 melibatkan 3 atribut dalam pembentukan model rule. Atribut tersebut yaitu lorong, lurus, dan kriminal. apabila lorong merupakan lorong utama maka kamera tipe B. Bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut kriminal. bila kriminal sering maka kamera tipe B. Sebaliknya bila jalan bukan rawan kriminal maka kamera tipe A. Pada rule ke-8 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-8, dapat dilihat pada flowchart Gambar 4.8.



Gambar 4. 8 Rule Ke-8 Dataset Kamera

Rule 9

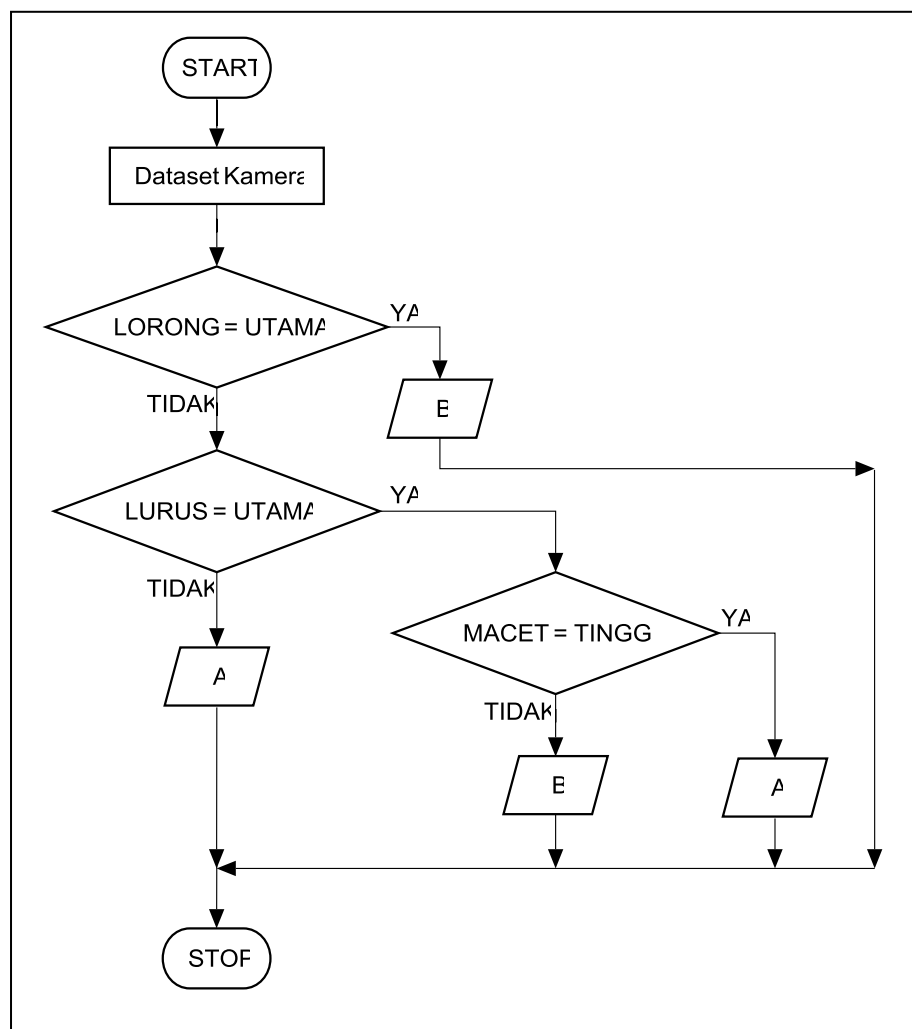
Tree ke-9 hasil simulasi algoritma decision tree C4.5 melibatkan 3 atribut dalam pembentukan model rule. Atribut tersebut yaitu lorong, lurus, dan kriminal. apabila lorong merupakan lorong utama maka kamera tipe B. Bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut kriminal. bila kriminal sering maka kamera tipe B. Sebaliknya bila jalan bukan rawan kriminal maka kamera tipe A. Pada rule ke-9 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-9, dapat dilihat pada flowchart Gambar 4.9.



Gambar 4. 9 Rule Ke-9 Dataset Kamera

Rule 10

Tree ke-10 hasil simulasi algoritma *decision tree* C4.5 melibatkan 3 atribut dalam pembuatan model rule. Atribut tersebut yaitu lorong, lurus, dan macet. Apabila lorong merupakan lorong utama maka kamera tipe B, bila tidak maka atribut selanjutnya yaitu atribut lurus dipersyaratkan. Bila bukan jalan lurus utama maka kamera tipe A. Jika lurus merupakan utama maka selanjutnya dipersyaratkan atribut macet. Bila macet tinggi maka kamera tipe A, sebaliknya bila bukan jalan macet maka kamera tipe B. Pada rule ke-10 tidak dihasilkan kamera tipe C. Untuk lebih mudah memahami rule ke-10, dapat dilihat pada flow chart Gambar 4.10.



Gambar 4. 10 Rule Ke-10 Dataset Kamera

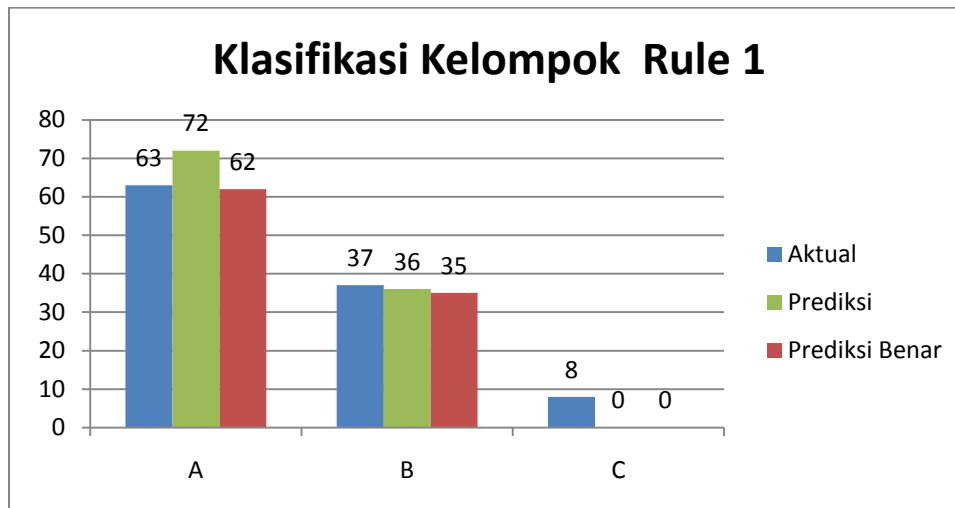
Penelitian ini akan membandingkan rule yang dibuat menggunakan algoritma C4.5 dengan rule yang dibuat oleh ahli. Dengan menggunakan rule tersebut, kemudian diberikan data testing untuk divalidasi menghasilkan nilai prediksi kelas kamera. Rule ke-1 menghasilkan klasifikasi benar sejumlah 97 data. Rule ke-2 menghasilkan klasifikasi benar sejumlah 97 data. Jumlah klasifikasi yang benar sesuai dengan rule ahli dapat dari 10 rule diatas dapat dilihat pada Tabel 4.3.

Tabel 4. 4 Klasifikasi Benar 10 Rule Decision Tree

Rule	Kelas A	Kelas B	Kelas C	Jumlah Klasifikasi Benar	Prosentase Klasifikasi Benar (%)
1	72	36	0	97	89,81
2	72	36	0	97	89,81
3	70	38	0	97	89,81
4	72	36	0	97	89,81
5	70	38	0	97	89,81
6	67	41	0	96	88,89
7	72	36	0	97	89,81
8	70	38	0	97	89,81
9	70	38	0	97	89,81
10	72	36	0	97	89,81

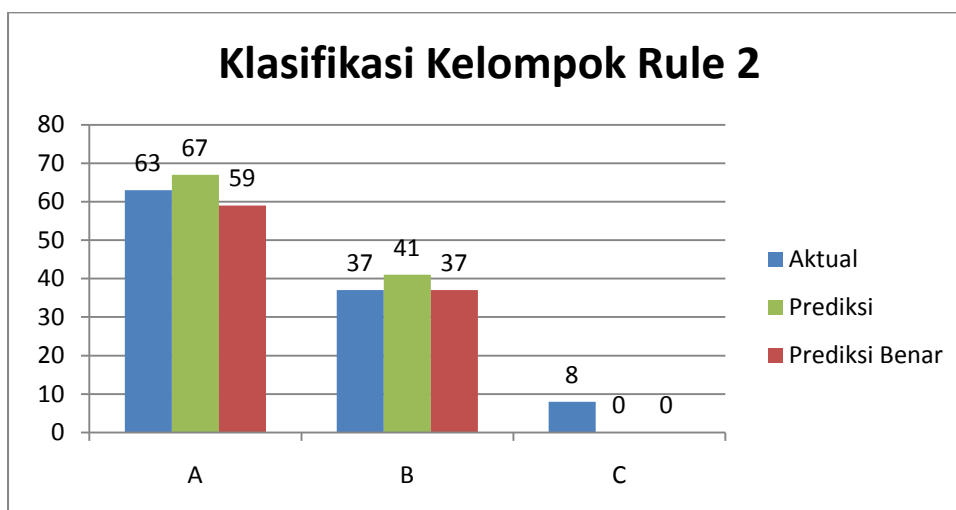
Rule-rule yang terbentuk dari algoritma decision tree C4.5 sebanyak 10 rule. Namun bila dicermati lebih jauh, maka sebenarnya hanya ada 3 kelompok rule yang terbentuk. Pertama, kelompok rule 1,2,4,7, dan 10. Kedua, kelompok rule 6. Dan ketiga, kelompok rule 3,5,8,dan 9. Kelompok pertama terdiri dari atribut lorong, lurus, dan macet. Kelompok kedua terdiri dari atribut lorong dan lurus. Dan kelompok ketiga terdiri dari atribut lorong, lurus, dan kriminal. Karena rule ke-1, 2, 4, 7, dan 10 sama maka prosentase klasifikasi benar bernilai sama yaitu 89,81%. Pada kelompok rule ke-1, dari hasil validasi terdapat jumlah kelas A sebanyak 72 data, dengan prediksi benar sejumlah 62 data. Pada kelas B

terdapat 36 data, dengan prediksi benar sebanyak 35 data. Dan kelas C bernilai 0. Nilai prediksi dan prediksi benar pada kelompok rule ke-1 dapat dilihat pada Gambar 4.11.



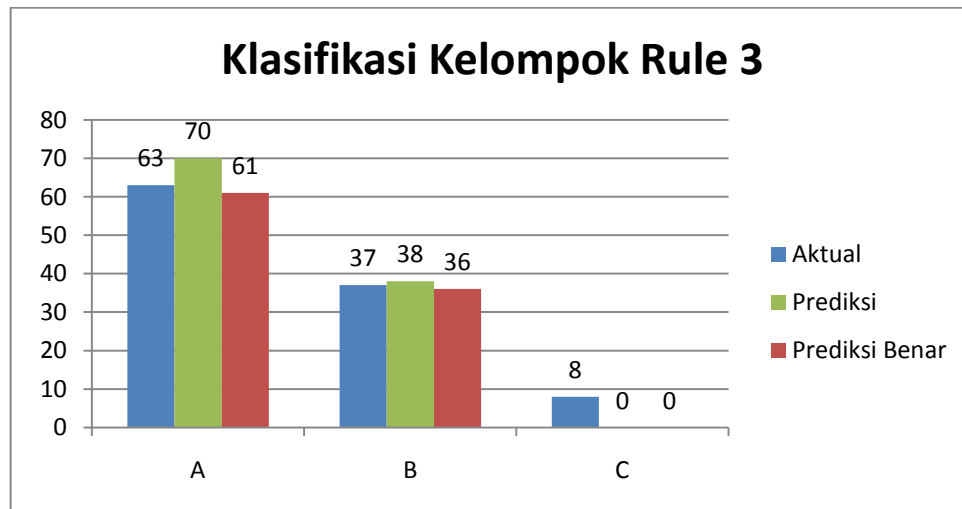
Gambar 4. 11 Klasifikasi benar Kelompok Rule ke-1

Pada kelompok rule ke-2, dari hasil validasi terdapat jumlah kelas A sebanyak 67 data, dengan prediksi benar sejumlah 59 data. Pada kelas B terdapat 41 data, dengan prediksi benar sebanyak 37 data. Dan kelas C bernilai 0. Nilai prediksi dan prediksi benar pada kelompok rule ke-2 dapat dilihat pada Gambar 4.12.



Gambar 4. 12 Klasifikasi benar Kelompok Rule ke-2

Pada kelompok rule ke-3, dari hasil validasi terdapat jumlah kelas A sebanyak 70 data, dengan prediksi benar sejumlah 61 data. Pada kelas B terdapat 38 data, dengan prediksi benar sebanyak 36 data. Dan kelas C bernilai 0. Nilai prediksi dan prediksi benar pada kelompok rule ke-3 dapat dilihat pada Gambar 4.13.



Gambar 4. 13 Klasifikasi benar Kelompok Rule ke- 3

Dari 9 input atribut dataset kamera, hanya 4 atribut yang digunakan dalam pembentukan model rule. Atribut tersebut yaitu lorong, lurus, macet, dan kriminal. Dari ketiga kelompok rule tersebut tampak bahwa hasil klasifikasi tidak ada yang masuk ke kelas C. Hal ini bisa disebabkan karena kelas C tidak berelasi kuat dengan salah satu atribut atau bisa disebabkan juga karena jumlah peluang munculnya kelas C yang sedikit sehingga dalam perhitungan dengan algoritma decision *tree* kelas C seperti ditiadakan. Pada beberapa penelitian hal ini disebut dengan *imbalanced dataset*.

Jika pembahasan kembali pada spesifikasi atau fitur kamera CCTV yang digunakan, maka akan jelas terlihat bahwa kamera tipe A dan tipe C memiliki fitur yang hampir sama. Satu-satunya perbedaan yang terlihat hanya pada kedalaman piksel, pada tipe A dapat merekam gambar dengan kualitas 704 x 480 piksel sedangkan pada tipe C 1280 x 1024 piksel.

Atribut lorong berelasi kuat dengan kelas B, sehingga apabila jenis jalan merupakan lorong utama maka langsung diputuskan menggunakan kamera tipe B. Namun beberapa titik lokasi yang tidak lorong utama pun oleh ahli masih diberikan kamera tipe B. Kamera tipe B memiliki perbedaan utama dengan kamera tipe A dan C pada kemampuan merekam gambar dalam kondisi gelap. Sehingga ahli memutuskan untuk memberikan kamera tipe B pada daerah gelap, misalnya jalan berupa lorong. Selain berelasi dengan atribut lorong, kamera tipe B juga berelasi dengan kriminal. Apabila lokasi pemasangan merupakan tempat rawan kriminal maka diberikan tipe B. Para ahli memberikan kamera tipe B dilokasi ini agar dalam keadaan gelap tetap dapat mengawasi jalan raya.

Atribut selain lorong utama maka bisa menggunakan kamera tipe A dan C. Namun pada kenyataannya lebih banyak dipasang kamera tipe A. Mungkin diperlukan atribut lain untuk membedakan penggunaan kamera tipe A dan tipe C, misalnya atribut harga. Karena kamera tipe A, B, dan C memiliki harga yang berbeda.

Dari ketiga model rule yang dihasilkan, tampak bahwa apabila lokasi bukan merupakan jalan lorong dan jalan lurus maka masuk pada kelas A. Pada kelompok rule 1 ditambahkan atribut macet sebagai persyaratan kelas A. Pada kelompok rule 3 ditambahkan atribut kriminal sebagai persyaratan kelas A.

Kamera tipe C tidak menjadi hasil dari ketiga model rule, hal ini bisa disebabkan 2 hal, pertama tidak adanya perbedaan yang nyata pada spesifikasi kamera tipe A dan tipe C, kedua jumlah penggunaan kamera C yang sedikit sehingga cenderung diabaikan atau dianggap salah pemasangan,

BAB V

KESIMPULAN

1. Klasifikasi dataset kamera CCTV dengan algoritma decision *tree* C4.5 didapatkan nilai akurasi sebesar 89,81%. Dengan pemilihan kamera CCTV yang tepat sesuai kondisi maka monitoring jalan raya dapat dilakukan secara optimal sehingga tindak kecelakaan, kriminalitas dan kecelakaan dapat dihindari.
2. Metode 10-fold validaton menghasilkan 10 model rule. Model rule yang sesuai untuk dataset kamera yaitu model rule ke-1, 2, 3, 4, 5, 7, 8, 9, dan 10. Model rule tersebut memberikan nilai akurasi klasifikasi yang sama besar yaitu 89,81%.

Halaman ini sengaja dikosongkan

DAFTAR PUSTAKA

1. Birch, I., Vernon, W., Burrow, G., dan Walker, J., (2014), "The Effect Of Frame Rate On The Ability Of Experienced Gait Analysts To Identify Characteristics Of Gait From Closed Circuit Television Footage." *Journal of Forensic and Legal Medicine* 54, no. 2, hal 159-163.
2. Bolbol, A., Cheng, T., Tsapakis, I., dan Haworth, J., (2012), "Inferring Hybrid Transportation Modes From Sparse GPS Data Using A Moving Window SVM Classification." *Computers, Environment and Urban Systems* 36, no. 6. hal 526-537.
3. Conche, F., dan Tight, M., (2006), "Use Of CCTV To Determine Road Accident Factors In Urban Areas", *Accident Analysis & Prevention*, hal 1197–1207.
4. Damjanovski, V., (2014), "CCTV Cameras", *Chap. 5 in CCTV*, 3rd edition, Elsevier Inc., hal 152-211.
5. Edelman, G, dan Bijhold, J., (2010), "Tracking People And Cars Using 3D Modeling And CCTV." *Forensic Science International*, hal 26–35.
6. Harrison, C., dan Donnelly, I.A., (2011), "A Theory Of Smart Cities." *Proceedings of the 55th Annual Meeting of the ISSS – 2011*, Hull.
7. HSSIN, B., MERBOUHA, A., EZZIKOURI, H., dan ERRITALI, M., (2009), "A Comparative Study Of Decision Tree ID3 And C4.5." (*IJACSA International Journal of Advanced Computer Science and Applications*).
8. Kurdi, H.A., (2013), "Survey on Traffic Control using Closed Circuit Television (CCTV)." *The International Conference on Digital Information Processing, E-Business and Cloud Computing (DIPECC2013)*, hal 133-138.
9. Li, T., Ma, S., dan Ogihara, M., (2004), "Entropy-Based Criterion In Categorical Clustering." *ICML '04 Proceedings of the twenty-first international conference on Machine learning*. 68.
10. Maa, X., Wu, Y., Wang, Y., Chen, F., dan Liu, J., "Mining Smart Card Data For Transit Riders' Travel Patterns." *Transportation Research Part C: Emerging Technologies* 36, hal 1-12.

11. Naphade, M., Banavar, G., Harrison, C., Paraszczak, J., dan Morris, R., (2011)"Smarter Cities And Their Innovation Challenges", *IEEE Journals & Magazines 44*, no. 6, hal 32-39.
12. Nerngchamnong, K., Kaviya, S., Fujii, Y., and Yupapin, P.P, (2011), "World Heritage City Surveillance System by a Smart CCTV System", *The 2nd International Science, Social Science, Engineering and Energy Conference 2010*. hal 321-327.
13. Park, H. H., Oh, G.S., dan Paek S. Y., (2012), "Measuring The Crime Displacement And Diffusion Of Benefit Effects Of Open-street CCTV In South Korea." *International Journal of Law, Crime and Justice 40*, no. 3 , hal 179-191.
14. Prashyanusorn, V., Prashyanusorn, P., Kaviya, S., Fujii, Y., dan Yupapin, P.P., (2011), "The Use of Security Cameras with Privacy Protecting Ability." *The 2nd International Science, Social Science, Engineering and Energy Conference 2010 (I-SEEC 2010)*. hal 301-307.
15. Raileanu, L.E., dan Stoffel, K., (2004), "Theoretical Comparison between the Gini Index and Information Gain Criteria." *Annals of Mathematics and Artificial Intelligence, 2004*, hal 77.
16. Seufert, M., Griepentrog, T., Burger, V., dan Hoßfeld, T., (2016), "A Simple WiFi Hotspot Model for Cities." *IEEE Communications Letters 20*, no. 2, hal 384 - 387.
17. Tan, P.N., Steinbach, M., dan Kumar, V., (2006), "Introduction to Data Mining", Pearson Education, Inc,
18. Winters, J. V., (2011), "Why Are Smart Cities Growing? Who Moves And Who Stays." *Journal of Regional Science 51*, no. 2, hal 253-270.
19. Wong, T. W., Tao, C. H., Cheng, Y. K., Wong, K. H., dan Tam, C. N., (2014), "Application of cross-ratio in traffic accident reconstruction." *Forensic Science International*, hal 19-23.
20. Desember 2016. https://en.wikipedia.org/wiki/Closed-circuit_television.
21. Trinity CCTV Solutions. Desember 2016. <http://trinitycctv.co.nz/cctv/learn-about-cctv/comparing-hd-and-sd-cctv-footage/>.

22. —. Desember 2016. <http://wireless-cctv.co.za/cctv-lenses-examples-comparisons/attachment/cctv-camera-lens-angles/>.
23. Wireless CCTV. Desember 2016. <http://wireless-cctv.co.za/cctv-lenses-examples-comparisons/attachment/cctv-camera-lens-images/>.
24. Discount Security Cameras. Desember 2016. <http://www.discount-security-cameras.net/cctv-video-resolutions.aspx>.

BIOGRAFI PENULIS



Arif Pribadi, anak kedua dari dua bersaudara. Merupakan anak kedua pasangan bapak Amat Alfian (Alm) dan ibu Sri Iriani. Sudah menikah dengan Endra Aryani dan memiliki dua anak, yaitu Aqila Pratama Gumilang dan Zafar Gigih Prawira. Lahir di Cilacap, Jawa Tengah. Dari kecil hingga menamatkan SMA berada di kota Cilacap. Sekolah dasar di SDN Sidanegara 09 Cilacap, sekolah menengah pertama di SLTP N 1 Cilacap, dan sekolah menengah atas di SMU N 1 Cilacap. Kuliah Strata 1 di kota Surabaya dari tahun 2004 sampai dengan 2008. Menjadi Pegawai Negeri Sipil dari tahun 2009 sampai dengan sekarang di Dinas Sosial Tenaga Kerja dan Transmigrasi Kabupaten Pekalongan, Jawa Tengah. Tahun 2015 mengikuti program beasiswa pendidikan S2 dari Kementerian Komunikasi dan Informasi Republik Indonesia bidang Telematika CIO (Chief Information Officer) di Teknik Elektro ITS Surabaya. Dapat dihubungi melalui e-mail: arifpribadi.1986@gmail.com.